Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Performance Changes in Eyelid Conditioning as Related to the Motivational and Reinforcing Properties of the UCS: M. A. TRAPOLD AND K. W. SPENCE 209
Absolute Judgments of Odor Quality: T. ENGEN AND C. PFAFFMANN 214
Order of Consideration of Different Types of Concepts: E. B. HUNT AND C. I. HOVLAND 220
Reinforcers and Reinforcement: Their Relation to Maze Performance: W. BEVAN AND R. ADAMSON 226
Relationship between MAS Scores and Association Values of Nonsense Syllables: E. P. TRAPP AND D. H. KAUSLER 233
Latency of Instrumental Responses as a Function of Compatibility with the Meaning of Eliciting Verbal Signs: A. K. SOLARZ 239
Effect of Discrimination Training on Auditory Generalization: H. M. JENKINS AND R. H. HARRISON 246
Effects of Extinction Trials on Discrimination Reversal: M. R. D'AMATO AND H. JAGODA 254
Learning and Performance as Functions of Ration Size, Hours of Privation, and Effort Requirement: G. D. JENSEN 261
Stimulus Meaning in Stimulus Predifferentiation: S. M. PFAFFLIN 269
Supplementary Report: Prompting Versus Confirmation in Paired-Associate Learning: J. O. COOK AND M. E. SPITZER 275
Supplementary Report: Context Effects on Absolute Judgments of Length: C.-A. MILLER AND T. ENGEN 276
Supplementary Report: Runway Performance as a Function of Reinforcement Schedule and Alley Length: E. L. WIKE AND R. REMPLE 277
Supplementary Report: Influence of Intertrial Interval during Extinction on Spontaneous Recovery of Conditioned Eyelid Responses: E. Y. BEEMAN, T. F. HARTMAN, AND D. A. GRANT 279

American Psychological Association

Vol. 59, No. 4 April 1960

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1960 the publication cost is $30 per printed page.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
BEATRICE B. HAMLIN, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1960 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 59, NO. 4 APRIL 1960

PERFORMANCE CHANGES IN EYELID CONDITIONING AS RELATED TO THE MOTIVATIONAL AND REINFORCING PROPERTIES OF THE UCS[1]

M. A. TRAPOLD AND K. W. SPENCE

State University of Iowa

In a recent series of lectures (Spence, 1956) the possibility was suggested that classical aversive conditioning is governed by a reinforcement principle. According to this view the development of the associative factor, habit strength (H), of a classically conditioned defense response is a function of the occurrence and intensity of the reinforcing UCS. This assumption, taken in conjunction with the existing assumption that drive level (D) is also a function of the intensity of the UCS, implies that the differences in performance level obtained with different intensities of the UCS reflect differences in both H and D. The problem of testing this theory thus requires that some means be found to separate out these two hypothetical functions of the UCS.

Since these lectures, several experimental studies, in which it is believed this separation was achieved, have reported findings that support this reinforcement interpretation of classical aversive conditioning (Spence, Haggard, & Ross, 1958a, 1958b). These experiments have involved holding the drive effects of the UCS constant while varying its reinforcing effects. In order to understand the logic of these experiments, as well as the one to be presented here, it is necessary to consider the source of motivation or drive in classical aversive conditioning. Since a CR occurs before the application of the UCS, we cannot appeal to the UCS on that particular trial as the source of drive. Furthermore, it has been shown (Spence & Runquist, 1958) that the drive cannot arise from emotional (fear) responses conditioned to the CS, since the latency of such fear responses is much longer than the 300–500-msec. latency of the conditioned eyeblink. Hence, the assumption has been that the drive level (D) is determined by previous presentations of the UCS. More specifically, the UCS is assumed to arouse a hypothetical emotional response (r_e), the effects of which are assumed to persist at least until the following trial. There is also the possibility that these emotional responses may become conditioned to other cues of the general stimulus situation. In any event, the assumption is made that the level of this hypothetical emotional activity is fairly constant from trial to trial, and that this level is, in large part, determined by the strength of the UCS. Drive level (D)

[1] This study was carried out as part of a project concerned with the influence of motivation on performance in learning under Contract N9 onr-93802, Project NR 154–107 between the State University of Iowa and the Office of Naval Research. Acknowledgement is made to John Hunter and Robert Weyant who served as research assistants in the above investigation.

is, in turn, assumed to be directly related to this level of emotional activity.

One of the implications of this analysis of the source of motivation in eyelid conditioning is that any application of a noxious stimulus in the situation will increase the level of r_e, and therefore the level of D. The above mentioned experiments capitalized on this implication by providing a technique by which D could be held constant for two groups of *S*s, and yet allowed the strength of the reinforcing UCS to be varied. Briefly, the technique consisted of using two strengths of air puff with each *S*. In the high reinforcement condition, the stronger air puff was always paired with the CS as a typical UCS, while the weak air puff was presented equally often in such a way that conditioning on these trials was precluded. In the low reinforcement group the reverse conditions held, the weak puff being paired with the CS and the strong puff always being presented so as to preclude learning. This procedure thus equated the average puff strength for both groups, and hence their drive level, but allowed the strength of the reinforcing UCS to be varied. Making use of the basic equation that performance (R) is proportional to $E = H \times D$, then, to the extent that D is equated for the two groups by this technique, any differences in their performance must be attributed to differences in H.

The results of these experiments have shown that when drive is so equated the asymptotic level of performance varies directly with the strength of the reinforcing UCS. These results have been interpreted as supporting the notion that the growth of H is a function of the strength of the reinforcing stimulus, and hence that a reinforcement principle operates in classical aversive conditioning. These experiments have also confirmed the assumption that D is a function of the intensity of the noxious stimulus employed, and that the unpaired puff is effective in manipulating the level of drive. Thus a group conditioned with a weak puff paired with the CS and a strong puff presented alone performed at a higher level than a group receiving a weak puff on both paired and unpaired trials.

One of the properties that has been ascribed to habit strength is that once accrued, habit strength is permanent, and cannot directly be reduced. In the 1943 version of his theory, Hull explicitly assumed that habit strength was a relatively permanent organization left within the nervous system of the organism. This assumption is also implicit in a more recent formulation (Hull, 1951). If this assumption about habit strength is taken in conjunction with the analysis of the role of the UCS presented in the above paragraphs, some interesting predictions concerning the effects upon performance of shifting the strength of the reinforcing UCS become evident.

(*a*) For a group, conditioned to the limit of learning with a weak air puff paired with the CS and a strong puff presented unpaired, and then switched to the reverse conditions, i.e., the strong puff paired and the weak puff unpaired, performance should show an increase following the shift, eventually reaching the level of a group trained under the high reinforcement condition from the beginning of training. The increase in the strength of the reinforcing UCS raises the level to which habit can grow, and the increase of habit to this new level is reflected in a performance increase.

(*b*) If a group, conditioned to the limit of learning with the strong puff paired with the CS and the weak puff presented alone is switched to the condition where the weak puff is now paired and the strong unpaired, no performance changes should be exhibited following the shift. Since habit strength is assumed not to decrease once accrued, and since the same drive level is maintained, there should be no change in performance.

(*c*) If a group, conditioned to the limit of learning with the strong puff as the reinforcing UCS, and the weak puff unpaired, is then switched to the condition in which both the paired and the unpaired puff are weak, we would expect a decrease in performance. This implication follows from the analysis of the source of motiva-

tion in aversive conditioning. In this instance the switch decreases the average puff strength, which would lead to a decrease in the level of hypothetical emotional activity and hence the level of *D*. To the extent that Prediction *b* is confirmed, performance changes in the present shift would reflect only a decrease in drive level. However, if the switch outlined in Prediction *b* does lead to a performance decrease, then the present group would reflect both a decrease in drive and a decrease in the reinforcing UCS.

The experiment reported here was conducted to provide data on these three predictions.

Method

Subjects. A total of 64 men from an introductory course in psychology was employed. Of these, five were eliminated for having met the criterion previously determined as defining a voluntary responder (Spence, 1953). The data of four other *S*s were also discarded, two for giving CRs on the pretest trials, and two who blinked continuously, making scoring of their records impossible. The remaining 55 *S*s provided three groups, one of which had 15, and two of which contained 20 *S*s.

Apparatus and method of recording. The *S* was seated in a dental chair in a semidarkened soundproof room. The *E* was in an adjoining room in which the stimulus controls and recording equipment were located. The apparatus for presenting the stimuli and recording the eyeblinks was identical with that used in previous studies from this laboratory (cf. Spence, 1953).

The CS consisted of the increase in brightness of a 6-cm. circular milkglass disk from a level of less than .004 mL. to 2.95 mL. The duration of the CS was 550 msec., with the UCS occurring 500 msec. after the onset of the CS. The duration of the UCS, an air puff of either .33 or 2.0 psi, applied to the right eye through a .062 in. diameter orifice, was limited to 50 msec. by a 100-v., 60-cycle ac-operated solenoid valve controlled by an electronic timer.

The word "ready," delivered by *E* over an intercom system, preceded each trial by 2, 3, or 4 sec., according to a prearranged schedule. The *S*s were instructed to blink upon presentation of the ready signal, and then fixate the circular disk in front of them. A CR was recorded whenever the record showed a deflection of 1 mm. or more in the range 200–500 msec. following the onset of the CS. Responses in the interval 0–200 msec., which were infrequent, were classified as original responses and not included in the data.

Conditioning procedure. Following the reading of the instructions, each *S* was given three presentations of the CS alone followed by a single presentation of the 2.0-lb. UCS alone. Immediately following these preliminary trials, conditioning was begun. Each *S* received a total of 260 trials, 130 conditioning trials which involved the paired presentation of the CS and UCS, and 130 which involved only the UCS following the ready signal. The order of presentation of these paired and unpaired trials was prearranged according to an irregular order in which the number of each was equalized in blocks of four trials. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec. and arranged according to a fixed schedule, were used throughout the experiment. Immediately following Trials 80 and 160, the experiment was interrupted while *E* checked the adjustment of the apparatus. No additional lights were turned on in the experimental room during this check, and it rarely took longer than one minute.

Experimental design. The experiment consisted of two phases. In Phase I, which lasted for 180 trials (90 conditioning trials), Groups A and B ($N = 20$ each) received identical conditions. For these groups the CS was always paired with the strong puff, the weak puff always being presented alone. For Group C ($N = 15$) the reverse conditions held, the weak puff always being paired with the CS and the strong puff always being presented alone.

Immediately following Trial 180, with no break in the continuity of the experiment, Phase II was begun. In this phase, Group A was switched to the reverse of its conditions in Phase I, the weak puff now being paired with the CS, and the strong puff being presented alone. Similarly, Group C was also switched to the reverse of its Phase I conditions, with the strong puff now paired with the CS and the weak puff unpaired. Group B was switched to receiving the weak puff on both the paired and the unpaired trials.

Results

Figure 1 presents the conditioning curves for the three groups. Here the

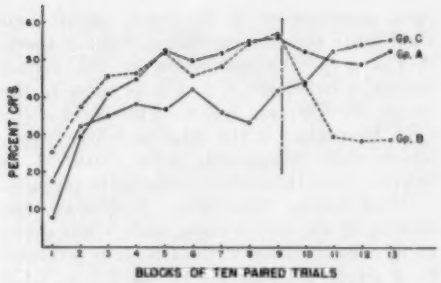


FIG. 1. Acquisition curves showing the percentage of CRs in successive blocks of 10 paired conditioning trials. The vertical line shows the point at which the shift was made.

percentage frequency of CRs is plotted against the number of conditioning trials (paired CS and UCS), which, of course, were the only trials on which CRs could be given. The vertical line represents the point at which conditions were switched from Phase I to Phase II. Groups A and B, which had identical treatment up to this switch point, can be seen to have attained almost identical performance levels. In contrast, Group C, which had the weak puff as the reinforcing UCS, reached an appreciably lower performance level than those of Groups A and B. Evaluation of this difference over the last 20 preshift conditioning trials revealed that the difference is statistically reliable ($P = .026$).[2]

Following the shift at Trial 90, it can be seen that Group B dropped quite rapidly down to a new level of performance. Comparing the last 20 trials before the shift (71–90) with the last 20 trials after the shift (111–130), reveals that this decrease is highly reliable ($P < .01$). Conversely, Group C increased following the shift, and came to approximate the preshift performance levels of Groups A and B. Again comparison of the last 20 preshift trials with the last 20 postshift trials reveals that this increase is significant beyond the 1% level. In contrast to Groups B and C, Group A maintained a relatively stable level of performance following the shift. Although Group A did show a slight decrease over Trials 90–120, the maximum amount of this decrease was only 7%, which is easily attributable to chance ($P = .29$).

[2] Because the distribution for frequency of response deviated considerably from normality, Mann-Whitney U tests were used to evaluate all between-group comparisons. Distributions of difference scores, however, were approximately normal, and hence t tests for related measures were used in all within-group comparisons. All reported probabilities are for tests of a two-tailed hypothesis.

DISCUSSION

The results of the preshift performances replicate those of earlier experiments in showing that when precautions are taken to insure equalization of drive level (D), performance is an increasing function of the strength of the reinforcing UCS. This is in line with the reinforcement position with respect to classical defense conditioning as outlined in the introduction. The postshift results also clearly support the predictions made by the reinforcement interpretation of this type of learning. When the strength of the reinforcing UCS was increased for Group C, hence raising the level to which H could grow, this group showed a corresponding increase in performance to the level of the previous high-reinforcement groups. The effect of the shift in strength of the reinforcing UCS on Group A, however, was quite different. In accord with prediction, habit strength being permanent, the decrease in the reinforcing UCS had little or no effect on performance. Taken together the results of these two groups lend considerable support to the notion that habit strength is an increasing function of the strength of the reinforcing UCS, and to the assumption that habit strength is a permanent acquisition that does not decrease when conditions are

changed in the direction of being less favorable for habit *growth*.

Taken in conjunction with the post-shift results of Group A, the decrease in performance following the shift for Group B can be said to represent the effects of decreased drive only. Since Group A showed no effect from merely decreasing the reinforcing UCS, the drop in Group B's performance must be due to the lower over-all puff strength, which, according to this analysis, is a primary determinant of drive level.

Summary

This study was concerned with several implications of a theory that assumes habit strength (H) to be a function of the intensity of the UCS in classical aversive conditioning. Groups A and B were conditioned with a strong air puff as the reinforcing UCS, a weak puff being presented equally often unpaired with the CS. Group A was then switched to the weak puff as the paired UCS and the strong puff unpaired. Group B was switched to the weak puff on both the paired and unpaired trials. Group C was conditioned with the weak puff paired with the CS and the strong puff unpaired, and then was switched to the reverse condition, strong puff paired and weak puff unpaired.

The preshift performances of Groups A and B were significantly higher than the performance of Group C. Following the shift, Group C gradually increased in performance to the preshift level of Groups A and B. Group B decreased significantly, and Group A showed no change in performance.

These results were interpreted as indicating that H is a function of the reinforcing UCS, that the extra unpaired puffs are effective in manipulating D, and that once accrued, H is permanent and does not decrease when conditions less favorable to habit *growth* are introduced.

References

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Hull, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1951.

Spence, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.*, 1953, **42**, 57–63.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

Spence, K. W., Haggard, D. F., & Ross, L. E. UCS intensity and associative (habit) strength of the eyelid CR. *J. exp. Psychol.*, 1958, **55**, 404–411. (a)

Spence, K. W., Haggard, D. F., & Ross, L. E. Intrasubject conditioning as a function of the intensity of the unconditioned stimulus. *Science*, 1958, **138**, 774–775. (b)

Spence, K. W., & Runquist, W. N. Temporal effects of conditioned fear on the eyelid reflex. *J. exp. Psychol.*, 1958, **55**, 613–616.

(Received June 1, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

ABSOLUTE JUDGMENTS OF ODOR QUALITY[1]

TRYGG ENGEN AND CARL PFAFFMANN

Brown University

This paper is concerned with the recognition of odor qualities and the application of information theory to the evaluation of the human observer in this modality. Garner and Hake (1951) have shown that an information measure, the bit, gives an index of the maximum number of single stimulus categories that can be discriminated without error by the method of absolute judgment. The present experiments were designed to determine the greatest number of different qualities that could be identified correctly by a label with different kinds, intensities, and numbers of odorants.

METHOD

Judgments were obtained from five *Ss*, one man and four women. All were college students who were paid $1.00 per hour. They had served previously in olfactory experiments on intensity differences. The experiments reported here extended over a period of 12 months and were interspersed among several other experiments on threshold and scaling.

The experiments were performed in a well-ventilated and air-conditioned 5 × 7-ft. room at a temperature approximately 24°C. In the first part of the experiment *S* learned to identify the stimuli. He was asked to label each odorant with his own association to that odorant (avoiding the restriction of an arbitrary code). This was repeated for each sample of odorants used. After the labels had been chosen, they were checked a second time to eliminate duplications and to give *S* the opportunity to make changes.

The odorants were kept out of sight behind a shield. Only one odorant was presented at a time in random sequences, and *S* was given unlimited time to identify it with the previoulsly assigned label. Each odorant in a set was presented 24 times for a total of 70 judgments of 5 odorants to 1152 judgments for 48 odorants. A correction procedure was used such that the correct label was indicated by *E* after *S* made his judgment. After a period of about 12 min., *S* was given a rest period while one or two other *Ss* were run. This arrangement seemed optimum for control of fatigue and adaptation. It should be noted that with five *Ss* it was not possible to obtain a balanced presentation of conditions, but no evidence was obtained that this restriction on experimental design had any appreciable effect on the results.

Two sets of 36 odorants were used. They are listed in Table 1 as different and similar. The set of so-called different odorants was selected to sample a wide variety of odorants, e.g., fragrant, spicy, acid, burnt, etc., and the similar set was selected to sample only sweet-fruity odorants. The latter set was obtained largely from Givaudan-Delawanna, Inc. and were described in their catalogue (No. 12, 1952) as having sweet and fruity odors. The remaining odorants were obtained from Fritzsche Brothers, Inc. and Brown University chemistry stock room. A few odorants (e.g., phenylethyl alcohol) appear in both sets. When fewer than 36 odorants were used in an experiment, they were selected at random from among the 36 odorants. In Table 1 a "5" opposite the odorant means that this odorant was included in the sample of 5 odorants, "7" that it was included in a sample of 7, etc. A "16" opposite the odorant indicates that the odorant was one of the 16 odorants presented at three levels of intensity. When the experimental procedure required it, geometric concentration dilution series were prepared of the odorant with a pipette technique, using benzyl benzoate as the dilutent (Beck, Kruger, & Calabresi, 1954). The odorants were presented in 10 × 75-mm. Pyrex test tubes stopped with an aluminum-wrapped cork.

Experiment I was concerned with the effect of variations in intensity and type of odorant on the amount of information transmitted. In a preliminary experiment each member of an undergraduate class of 40 students had been presented one test tube from a dilution

[1] This research was supported by the General Foods Corporation. Part of the results was presented at the Thirtieth Annual Meeting of the Eastern Psychological Association in Atlantic City, April 4, 1959.

TABLE 1

CHEMICAL AGENTS USED IN THE EXPERIMENTS
(NUMBERS OPPOSITE AGENTS ARE EXPLAINED IN THE TEXT)

Different	Similar	YM
Acetanisol—16	Acetate C-8—24, 10	
Acetic acid—24, 15, 10, 7, 5	Acetate C-9—24, 10, 7, 5	
Acetone—24, 15	Acetate C-10—24, 7, 15	
Amyl acetate—24, 16, 15	Acetate C-11—24, 10	
Allyl Caproate—24, 16, 15, 10, 7	Aldehyde C-7	
Benzaldehyde, N.F.—16	Aldehyde C-8	
Benzene—24, 16, 10, 7	Aldehyde C-10—24, 10	
Benzyl Cinnamate—7, 5	Aldehyde C-14 (peach)—24, 15	
N. Butanol—24, 7	Aldehyde C-14 (pure)—24, 10	
N. Butyric acid—10	Aldehyde C-16—24, 15	
Camphor (lt. syn.)—24, 16	Aldehyde C-18—24, 10	
N. Caprylic	Allyl Caproate—24, 15, 7	
Citral—24, 15	Amyl acetate—24	
Clove oil—24, 15	Amyl Phenylacetate—24, 15, 10, 5	
Coumarin—24, 15	Amyl Propionate	
Diacetone alcohol—24	Benzyl Acetate Coeur—10	
Ethyl acetate—7	Benzyl Butyrate—24, 15, 10	
Eugenol—24, 16, 15, 10, 7	Cinnamyl acetate—24, 15, 10, 5	
Guaiacol—24, 16, 15	Cinnamyl Butyrate—24, 15	
Heptanol—24, 16, 15	Cinnamyl Propionate—24	
Heptane—24, 15, 7	Citral—24, 10	
Hexane, N.—24, 16, 10	Citronellyl acetate—24, 5	
Indol—24, 15, 10	Citronellyl Butyrate—24, 15, 7	
Isopropanol—16, 15, 10, 5	Dimethyl Anthranilate	
Linalool—24	Geranyl Butyrate	
Menthol—24, 16, 10	Geranyl Propionate	
Methanol—16, 10, 5	Linalool—24, 15, 7	
Methyl Salicylate—24, 16, 5	Menthol	
Musk—24, 16	Methyl Anthranilate Standard	
Novoviol, alpha—15	Methyl Salicylate—24, 15, 7	
Phenylethyl alcohol—24, 16, 15	Phenylethyl acetate—24, 15, 5	
Pyridine—24	Phenylethyl alcohol—24, 15	
Safrole—24	Phenylethyl Isovalerate—24, 15, 10, 7	
Vanillin—24	Raspbery aldehyde	
Butyl acetate—15	Tolyl acetate—24, 15	
Nitrobenzene	Vanillin	

series of amyl acetate and was asked whether the odorant was "weak," "medium," or "strong." It was learned that a 6.25% solution of amyl acetate was, on the average, judged to be of medium intensity or "strength." This standard solution was then cross-matched by the five *S*s used in the present experiments with the 23 other different odorants in Table 1. Ten such cross-matches were made by each *S* for each odorant, varying the order and range of concentration dilutions of the comparison odorant to prevent context effects. This experiment yielded 24 odorants for each *S* of subjectively equal and medium intensity but different in quality for the first series of judgments. The same 24 odorants, but each in a 100% solution, were presented for a second series of judgments. A third selection of odorants was made of 24 odorants of similar odor qualities, each being presented in 100% solutions, as in the second series above.

Experiment II evaluated the information transmitted when both qualities and intensities of each odorant varied. Sixteen different odorants were selected, each in three concentration dilution values, 100, 25, and 6.25%. All of these concentration dilution values were clearly above threshold and more than 1 jnd apart. An earlier experiment (Engen & Pfaffmann, 1959) had indicated that these *S*s could transmit at least 1.5 bits of information about odor intensity. The purpose of this experiment was to determine whether or not the information transmitted for intensity would combine with the information transmitted for quality.

Experiment III was concerned with "maxi-

TABLE 2

MEANS AND *SD*s OF BITS OF INFORMATION TRANSMITTED BY 5 *S*s FOR VARIOUS SAMPLES OF ODORANTS

Sample of Odorant	Mean	*SD*
Different odorants of equal and medium intensity	4.00	.17
Different odorants unequal intensities	4.03	.19
Similar odorants unequal intensities	3.86	.43

mum channel capacity" (Alluisi, 1957) for odor quality and the extent to which the amount of information transmitted in the experiment above may be limited by the number of stimuli presented for judgment. Thus 5, 7, 10, 15, 24 and 36 odorants in 100% solutions were presented consecutively to each *S* under the same conditions as above. Both similar and different odorants were used.

RESULTS

The amount of information transmitted by each *S* under each condition in Exp. I was determined by a method suggested by Garner and Hake (1951). The average amount of information for all *S*s is presented in Table 2. First, the results indicate that the average *S* could transmit about 4.00 bits of information about odor stimuli which represent a wide variety of qualities (i.e., different odorants). According to information theory, then, *S* could use about 16 categories (the antilog of 4.00) without error in identifying odors. Second, there is the suggestion that this result will not vary appreciably with the intensity of these odorants, i.e., the distinctiveness of the odorants did not seem to be impaired by matching them in subjective intensity, because the matched odorants of medium intensity yielded about the same result (4.00 bits) as when they were unmatched (4.03 bits). Third, the selection of qualities to be judged is an important factor, because when the odorants were all similar (sweet-fruity) only 3.86 bits of information (14 categories) were transmitted. The similar odorants also seem to yield a more variable performance, as indicated by *SD*s which give an index of individual differences.

It should be noted that results for unmatched intensities must be interpreted with caution, because the two variable aspects of the stimulus, quantity and intensity, are not orthogonal. The presentation of each odorant at 100% concentration dilution provides chemically specified qualities but (psychological) intensities of indeterminate values, thereby precluding equally frequent occurrences of given quality and intensity combinations as well as a more appropriate and complex analysis (McGill, 1954) of the data. This is a statistical problem separate from the problem of psychological additivity of quality and intensity.

Quality and intensity were orthogonal by design in Exp. II in which each of 16 odorants were presented at three levels of intensity. These multidimensional judgments were analysed according to McGill's multivariate procedure, and the result of this analysis is presented in Table 3. The table shows that 3.51 bits of

TABLE 3

MEANS AND *SD*s OF BITS OF INFORMATION TRANSMITTED BY FIVE *S*s FOR 3 LEVELS OF INTENSITY FOR 16 ODORANTS

Transmission	Mean	*SD*
Quality—Responses	3.51	.14
Intensity—Responses	.27	.05
Interaction	.09	.01

information were transmitted about quality and .27 bits about intensity. Taking into account the small interaction of .09 bits per stimulus, it is indicated that 3.60 bits were transmitted about quality for a specified intensity and .36 bits about intensity for a specified quality. It is evident again that intensity did not add much to the distinctiveness of the odor stimuli.

Figure 1, which presents the data from Exp. III in the dashed-line curve, shows that the data presented in Table 2 are not limited by the 24 categories presented for judgment. Performance improves up to 24 categories but no further improvement results from presenting a larger number of categories, while efficiency is greatly reduced. Moreover, Fig. 1 shows that the difference between different and similar odorants is verified. It is interesting to note that performance was never perfect even when fewer than 16 alternative stimuli were presented. According to these results there is a bias in the prediction of the number of categories *S* can use without or with a minimum of error. It does not, however, affect seriously the interpretations made of the results.

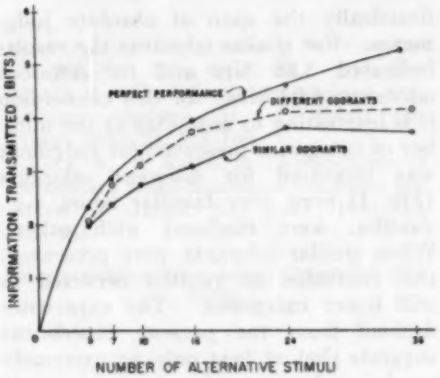


FIG. 1. Channel capacity for absolute judgments of odor qualities. Amount of information transmitted against number of alternative stimuli for two kinds of odors (described in text).

DISCUSSION

It is often assumed that the human *S* has a great capacity for discriminating and remembering odor qualities (Behnke, 1954). There has also been much discussion about the difference in the human *S*'s ability to remember and discriminate odor intensity and odor quality. Information theory provides one measure relevant to this problem since the bit permits the comparison of absolute judgments obtained under different experimental situations. Miller (1956) has already done this for several sense modalities.

While a relatively practiced *S* can transmit a maximum of 2.0 bits of information about odor intensity (Engen & Pfaffmann, 1959), the present results indicate that his channel capacity is twice as high for odor quality. Both sets of results appear to have some generality since they were obtained under a reasonably wide range of experimental conditions. As to the comparison of absolute judgments of odor intensity and quality, then, the informational analysis of absolute judgments probably indicates that the popular notion is in error. Given some practice, the human *S*'s ability to judge odor intensity approaches his ability to judge auditory and visual stimuli. His ability to identify odor quality is not phenomenal but falls within the range of his ability to identify multidimensional stimuli in other sense modalities (Miller, 1956).

One might raise the problem of the generality of the present results to other odor samples. It is practically impossible to sample all the odorants that could have been used in these experiments. However, while the present quality experiments indicate a significant variation in amount of transmitted information due to the selection of odorants (similar and different), this variation is not large enough to change

drastically the span of absolute judgments. For similar odorants the results indicated 3.86 bits and for different odorants, 4.03 bits. In this connection it is interesting to note that as the number of categories presented for judgment was increased for different odorants (Fig. 1) even very familiar odors, e.g., vanillin, were confused with others. When similar odorants were presented, this confusion of vanillin occurred in still fewer categories. The experience derived from the present experiment suggests that at best only an extremely careful selection of odorants would raise the channel capacity for quality under the present conditions.

The results also have some further bearing on the problem of multidimensional judgments. It was surprising, judging on the basis of similar experiments in other sense modalities (Beebe-Center, Rogers, & O'Connell, 1955; Eriksen, 1954; Klemmer & Frick, 1953; Pollack, 1953; Pollack & Ficks, 1954), that the presentation of three intensity levels of each of 16 qualities (i.e., the sum of transmitted information for quality and intensity) did not yield improvement over quality alone. It was expected on the basis of the results of these experiments that the performance would be better than that for either intensity or quality alone but would fall short of the sum of the two. Analysis of individual judgment matrices indicated that very few errors were made on quality, but the intensity levels were badly confused. These confusions are most likely due to the fact that the three levels of intensity for each odorant were neither subjectively matched nor spaced on an equal discriminability scale. They were only equal in concentration dilution with the result that the strongest concentration dilution of Odorant A could be subjectively equal to the lowest concentration dilution of Odorant B, etc. In combining dimensions in such experiments one should deal with psychophysical rather than physical values. This seems especially important in the present case, but may also explain why, as information is added to sensory displays, the accuracy for any particular variable is so often drastically reduced (Miller, 1956). Just as there is an optimum spacing (jnds) of stimuli in a unidimensional experiment (Garner, 1953) one may expect there is an optimal combination of physical values in a multidimensional experiment.

Summary

A series of experiments were carried out to determine how well the human *S* can identify odor qualities by label (S's own association to the odorant) with different kinds, intensities, and numbers of odorants. Informational analysis of these identifications yielded approximately 4.0 bits. The results were influenced by variations in the sample of odor qualities but not appreciably by intensity characteristics of the odor qualities. The information transmitted for a combination of 1.5 bits obtained for intensity earlier and the 4.0 bits obtained for quality did not exceed that for quality alone. The discussion of the results compared the span of absolute judgments in olfaction and other sense modalities with respect to unidimensional and multidimensional stimuli.

References

Alluisi, E. A. Conditions affecting the amount of information in absolute judgments. *Psychol. Rev.*, 1957, **64**, 97–103.

Beck, L. H., Kruger, L., & Calabresi, P. Observations on olfactory intensity: I. Training procedures, methods and data for two aliphatic homologous series. *Ann. NY Acad. Sci.*, 1954, **58**, 225–238.

Beebe-Center, J. G., Rogers, M. S., & O'Connell, D. N. Transmission of information about sucrose and saline solutions through the sense of taste. *J. Psychol.*, 1955, **39**, 157–160.

Behnke, A. R. Basic odor research correlation: Introduction. *Ann. NY Acad. Sci.*, 1954, **58**, 15–21.

Engen, T., & Pfaffmann, C. Absolute judgments of odor intensity. *J. exp. Psychol.*, 1959, **58**, 23–26.

Eriksen, C. W. Multidimensional stimulus differences and accuracy of discrimination. *USAF, WADC Tech. Rep.*, 1954, No. 54–165.

GARNER, W. R., & HAKE, H. S. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.

KLEMMER, E. T., & FRICK, F. C. Assimilation of information from dot and matrix patterns. *J. exp. Psychol.*, 1953, **45**, 15–19.

MCGILL, W. J. Multivariate information transmission. *Psychometrika*, 1954, **19**, 97–116.

MILLER, G. A. The magical number seven, plus or minus two: Some limitations on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81–97.

POLLACK, I. The information of elementary auditory displays. II. *J. Acoust. Soc. Amer.*, 1953, **25**, 765–769.

POLLACK, I., & FICKS, L. Information of elementary multi-dimensional auditory displays. *J. Acoust. Soc. Amer.*, 1954, **26**, 155–158.

(Early publication received December 29, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

ORDER OF CONSIDERATION OF DIFFERENT TYPES OF CONCEPTS[1]

EARL B. HUNT[2] AND CARL I. HOVLAND

Yale University

In the learning of some types of concepts individuals acquire the correct classification without any conscious or clear-cut plan of attack. This is typical of much of our perceptual learning and even of many simple conceptual tasks. But in other types of situations, for example those that are involved in scientific analysis, individuals characteristically go about the task of finding the concept by some type of deliberate strategy. This was the most common type of approach in the acquisition of concepts involving geometrical figures in the studies of Hovland and Weiss (1953) and of Bruner, Goodnow, and Austin (1956). These studies fostered deliberate approaches because of the tasks chosen and the nature of the instructions. For such problems one can logically distinguish a number of different kinds of concepts, derived from the types of classifications of instances described by the logical operations "and" and "or." In the *conjunctive* type all of the instances have features in common, so the concept is one where each instance possesses Characteristics A *and* B, or A *and* B *and* C, etc. In the case of *disjunctive* concepts, all instances have one *or* another feature (cf. Bruner, Goodnow, and Austin's [1956] example of the concept "strike" in baseball, where a strike is defined as either a pitched ball which crosses the plate between the batter's knees and his shoulders *or* is alternatively any pitch struck at by the batter which fails to be sent into the field). The authors just mentioned found this type of concept exceedingly difficult to learn. The third type of concept, discussed most extensively by Smoke (1935), is *relational*, in which the common properties are sets of relationships rather than common specific stimulus elements. For example, the concept of isosceles triangle involves the common relationship of equality between two of the sides, and the positive instances may have widely differing size and shape.

Subjects typically structure a concept problem in one or another of these ways and sometimes pursue their initial formulation until the evidence for its being unsuitable for the problem becomes overwhelming. In machine simulation of human concept learning it then becomes important to know the order in which solutions of each type will be tried (cf. Hovland, 1959). Since there may be individual differences in the order in which alternatives are considered it may also become necessary to establish the frequency of appearance of a particular order for different segments of the total population.

Order of attack may also be influenced by the perceptual characteristics of the stimuli to be classified. Within a single type of concept (conjunctive), variations in order of consideration have been demonstrated

[1] This research was financed by a Ford Foundation grant to Carl I. Hovland for the investigation of concept formation: this support is gratefully acknowledged.

[2] The present experiment was performed while this author held a General Electric Educational and Charitable Fund Fellowship at Yale University.

by Heidbreder and her associates (1945) and by Grant (1951). The *S*s in these experiments first tried to classify on the basis of form; when this was unsuccessful they tried color and, finally, number of figures. One might also anticipate an interaction between stimulus characteristics and type of concept selected such that when certain dimensions are used conjunctive concepts would be considered first whereas with others disjunctive or relational concepts would be tried first.

The present experiment was designed to obtain a systematic comparison of the frequency with which *S*s choose each of the three concept types as their initial basis for organizing instances. A series of geometrical forms labeled as positive or negative instances was presented to *S*. The instances were chosen in such a way that they could be organized with equal logical validity on a conjunctive, disjunctive, or relational basis. Which type of concept *S* utilized was determined on the basis of which of a series of test instances he selected as positive instances of the concept which he had derived from the training series. Each *S* was given one of three different series of positive and negative instances in which color, form, and number dimensions of the stimuli were rotated as appropriate bases for particular types of concept. This design made it possible to evaluate any interaction between stimulus characteristics and type of concept utilized.

METHOD

Stimuli. The stimuli were designs of the type illustrated in Fig. 1. The figures were in purple ditto ink; the stripes in the upper left and lower right corners were colored by crayon.

The stimuli involved six dimensions: color of the top stripe, color of the bottom stripe, number and form of the upper figure, number and form of the lower figure. The values used were: color (red, green, light brown, black); number (one through four); figure (triangle, cross, fleur-de-lis, and star). The sets of 15 stimuli each used during training were reproduced on 8½ x 11-in. sheets; two contained six instances each, the third three. Assignment of stimuli to a particular location and page was random. The 16 similar stimuli used in testing were presented eight to a page.

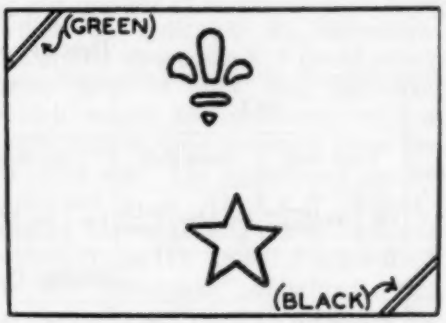


FIG. 1. Illustrative stimulus pattern.

Design. Each *S* was randomly assigned to one of three conditions corresponding to the set of stimuli to be used in training. During training each *S* received a set of positive or negative instances from which one conjunctive, disjunctive, or relational concept could be derived. The three training sets provided rotation of the dimensions used for each type of concept. Following the training set *S* selected concept exemplars from a test set consisting of unlabeled stimuli.

Each training set contained nine positive instances (labeled "Alpha") and six negative instances (labeled "not Alpha"). These completely defined one conjunctive, one relational, and one disjunctive concept, provided *S*'s hypotheses involved less than three dimensions.[3] The sets of instances used in each of the three training series are presented in Table 1. It will be seen that in Training Series I the conjunctive concept "upper red stripe and lower black stripe" would be consistent with this set of positive and negative instances. The relational concept "same number of upper and lower

[3] Examination of the pattern of selection of test instances by *S*s showed that *S*s did not use categorizing rules based on three or more dimensions.

TABLE 1

DESCRIPTION OF STIMULI[a]

	Set I						Set II						Set III					
	Upper Half			Lower Half			Upper Half			Lower Half			Upper Half			Lower Half		
	Fig.	No.	Stripe Color	Fig.	No.	Stripe Color	Fig.	No.	Stripe Color	Fig.	No.	Stripe Color	Fig.	No.	Stripe Color	Fig.	No.	Stripe Color
Training: Positive Instances																		
1.	C	4	R	F	4	Bk	C	2	G	C	3	Br	S	1	Bk	T	4	Bk
2.	T	4	R	F	4	Bk	C	2	Bk	C	3	Br	S	3	Bk	T	4	Bk
3.	C	1	R	C	1	Bk	S	2	G	S	3	R	S	1	R	T	1	R
4.	C	1	R	T	1	Bk	S	2	G	S	3	Br	S	1	R	T	4	R
5.	C	3	R	S	3	Bk	F	2	G	F	3	G	S	1	Br	T	3	Br
6.	C	2	R	S	2	Bk	T	2	G	T	3	Br	S	1	G	T	4	G
7.	C	2	R	T	2	Bk	T	2	G	T	3	Bk	S	1	G	T	3	G
8.	F	3	R	F	3	Bk	F	2	R	F	3	Br	S	2	Br	T	4	Br
9.	S	1	R	F	1	Bk	S	2	Br	S	3	Br	S	4	R	T	4	R
Training: Negative Instances																		
1.	T	1	R	S	4	G	S	2	Bk	C	4	Gr	S	2	R	S	1	Bk
2.	T	2	G	T	4	Bk	T	4	Bk	C	3	Bk	C	2	Gr	T	1	Bk
3.	S	2	R	T	3	Br	T	2	Br	F	2	Bk	S	3	Gr	F	2	Br
4.	S	1	Br	S	3	Bk	S	3	Br	F	3	G	F	3	R	T	2	Br
5.	F	3	R	C	2	G	F	2	R	S	1	R	S	4	Br	C	3	G
6.	F	4	Bk	C	1	Bk	C	1	R	T	3	R	T	4	Bk	T	3	R
Test Instances																		
A.	F	1	G	S	1	Bk												
B.	T	2	Br	S	3	G												
C.	S	2	Bk	T	2	Br												
D.	S	1	R	S	4	Bk												
E.	S	3	R	T	2	Bk												
F.	C	4	R	C	2	R												
G.	T	4	R	C	4	Br												
H.	S	3	Br	T	3	G	(Test instances same for all groups)											
I.	T	1	R	T	4	Bk												
J.	S	4	R	T	1	Bk												
K.	F	2	Bk	S	3	Bk												
L.	T	2	R	C	3	R												
M.	C	4	G	F	3	Br												
N.	F	3	Bk	F	1	Bk												
O.	C	2	G	F	1	Br												
P.	F	2	Bk	C	3	R												

[a] In the training set, positive and negative instances were interspersed randomly. The description code for the figures is S—star, C—cross, F—fleur-de-lis, T—triangle. For colors it is R—Red, G—Green, Br—Brown, Bk—Black.

figures" would be equally correct, as would the disjunctive concept "crosses in the top row and/or fleur-de-lis in the bottom row."

Following presentation of the training series *S*s were given the test set: 16 stimuli of the same type, but without labels. A description of these stimuli is given in the lower portion of Table 1. This set was the same for all *S*s. Subjects were instructed to pick out those instances in the test set which they considered to be positive instances of the concept they had derived from the training set. One set of four instances would be selected if *S* had derived a conjunctive con-

cept from the training series, another set of four if he had utilized a disjunctive one and a third set of four if it were relational. If the reader directs his attention to the relationship between the test set and Training Series I he will see that four of the patterns (D, E, I, J) are further instances of the conjunctive concept of this training series, four others (A, C, G, H) are positive instances of the relational concept, and four others (F, M, N, O) are positive instances by the disjunctive definition. In addition four "dummy" instances (B, K, L, P) were included which could not be considered as positive instances by any of the three categorizing rules derivable from the training series. Similar relationships between training and test stimuli obtained for the other two series.

The *Ss* could choose as many "positive" instances as they wished. If the majority of the test stimuli selected by *S* were consistent with the conjunctive definition he was categorized as having utilized a conjunctive concept; a similar categorization was applied for relational and disjunctive concepts. If the majority of *S*'s choices were of "dummy" instances or if he did not place a majority of his choices in one of the other three categories he was classified as "inconsistent."

Subjects. Thirty-seven students from the New Haven College day school and 25 from the night school served as paid *Ss*.[4] One *S* in each group had to be dropped for failure to understand the instructions. No *S* had previously participated in a concept formation experiment.

Procedure. Each *S* was handed an instruction sheet and an envelope containing the test and training series. The *E* read the instructions, answered specific questions, and described the six dimensions. The *S* then read the training sheets, returned them to the envelope, took out the test sheets, selected his choice of instances, and handed in the material. This procedure made it possible to present all training sets to the group and to randomly assign *Ss* to each training set.

Results

The *S*'s utilization of a particular concept type was evaluated in terms of the number of stimuli of each type which he indicated as exemplars. As stated above, each *S* could select from zero to four test instances which would be consistent with a particular concept derivable from the training set. The experiment can be analyzed as a $2 \times 3 \times 4$ "mixed" design corresponding to Lindquist's (1953) Type III. Each *S* is involved in one "treatment" of schools. one of "training set" and all measures for the learning of a particular concept type. In this analysis it is necessary to have proportional numbers of *Ss* in each of the three conditions (types of training sets) across schools. For this reason the number of cases for each training series within schools was randomly reduced to 11 (for the day school) and seven (for the night school). Analysis of variance methods were used although the scores within each treatment were not normally distributed and the range of scores was curtailed. This makes a conservative interpretation of significance levels mandatory. It will be seen that the differences presented in Table 2 are large. There is no evidence that the two school populations differed on the measures taken.

TABLE 2

Analysis of Variance of Choices of Instances on the Test Series

Source	*df*	*MS*	*F*
Between *Ss*	53		
Schools	1	.26	.20
Training sets	2	7.59	5.97*
Stimulus × Schools	2	.83	.65
Error (between)	48	1.27	
Within *Ss*	162		
Concepts	3	32.01	14.28*
(Relational and conjunctive) vs. (Disjunctive and error)	1	96.00	42.84*
Relational vs. conjunctive	1	.02	.01
Disjunctive vs. error	1	.02	.01
Concepts × Schools	3	.70	.31
Concepts × Training Sets	6	.74	.33
Concepts × Schools × Training Sets	6	3.82	1.70
Error (within)	144	2.24	

* $P < .01$.

[4] The authors wish to express their thanks to the administrative and teaching staff of New Haven College for their cooperation, and to Benton S. Underwood and Fred Sheffield for helpful suggestions for the writing of this report.

TABLE 3

MEAN NUMBER OF INSTANCES CHOSEN FROM THE TEST SERIES

Training Set	Type				Mean Total
	Relational	Conjunctive	Disjunctive	"Inconsistent"	
	Day School (N = 33)				
I	1.73	2.00	1.09	.91	5.73
II	1.46	1.82	.09	.00	3.37
III	1.82	1.27	.45	.36	3.90
	Night School (N = 21)				
I	3.43	1.71	.43	.29	5.86
II	1.43	1.29	.00	.14	2.86
III	1.14	2.56	.57	.86	5.14

Mean total by training set: I = 5.78, II = 3.17, III = 4.39.
Mean total by schools: Day = 4.33, Night = 4.57.

The mean numbers of instances selected by *S*s, by concept type and training set are presented in Table 3. For the three training sets there was a tendency to check different numbers of instances. Possibly this related to problem difficulty and the degree of certainty felt by *S*s after they had seen a particular training set. It will be remembered that an *S* who learns, with absolute certainty, one categorizing rule will check exactly four instances. The major part of the variance was due to *S*s concentrating their choices on relational and conjunctive concept exemplars much more frequently than any other type of instance in the test series.

The frequency of choice of particular types of concept exemplars did not vary with training set, as is shown by the nonsignificant Concepts × Training Set interaction.

The analysis of mean number of instances of each type of concept chosen by *S*s is supported by an examination of *S*s' concentration of choices of instances which are consistent with a particular type of concept derivable from the training set. Four *S*s concentrated their choices on disjunctive instances, eight on "dummy" instances (or had no majority), 24 concentrated on conjunctive choices and 24 concentrated on relational choices. There was no difference in frequency of choice concentrations between training sets or between the day- and night-school *S*s.

DISCUSSION

The above results provide a tentative answer to the question concerning the order in which different types of concepts are considered. Under our experimental conditions both conjunctive and relational concepts are selected significantly more frequently than disjunctive ones. Accordingly, we plan to program our initial computer simulations so that conjunctive possibilities are considered alternately with relational ones. Only when both have been explored without success will the program of the computer consider possible solutions involving disjunctive concepts. The present experiment does not provide information as to what determines an *S*'s initial choice of conjunctive or relational concept. This information will be needed to specify which of two alternative programs should be applied to simulate the behavior of a particular individual on a given concept learning problem.

A number of limitations on the generality of the present results must be considered. One concerns the population of concepts from which the limited set used in the present experiment was derived. Are our relational concepts, for example, typical of the population of such concepts or is there some tenddency for investigators to select ones which are particularly easy to learn? And are our results in which "sameness" was used for the relational concept equally applicable to relationships such as "subordinate to" or "bigger than"?

There may be some question as to whether the three types of concepts are distinctive psychologically. For example, some *S*s use *and* and *or* interchangeably, some describing a particular

class of stimuli as consisting of red *or* green triangles while others describe the same grouping as red *and* green triangles. Similarly, when pairs of large and small circles and triangles are the stimuli some derive the concept of "two large triangles *or* two small triangles" (a disjunctive concept) while others speak of a grouping in which "the triangles are both the same size" (a relational concept).

Another restriction on generality may arise from the use of college students in the present experiment. It is possible that the preference for conjunctive and relational concepts is a product of learning and that children may tend to select disjunctive concepts more frequently than adults.

No evidence was obtained of interaction between the perceptual characteristics of the stimuli and the type of concept chosen, although interaction effects had been anticipated. It is felt that further exploration of this problem is indicated with other quite different types of stimulus material, including some where the dimensions are more difficult to discriminate. But until evidence for interaction is obtained it does not appear necessary to provide alternative computer programs for different types of stimuli in simulation studies.

Summary

The *Ss* in a visual concept formation task could derive a correct concept from the instances presented on a conjunctive, disjunctive, or relational basis. It was found that conjunctive and relational solutions were utilized much more frequently than disjunctive ones. No significant interaction effects were obtained between the particular aspects of the stimulus involved (color, number, or type of figure) and the frequency of utilization of the different types of concept. The experiment is discussed in relation to machine simulation of human concept learning.

References

Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking.* New York: Wiley, 1956.

Grant, D. A. Perceptual versus analytic responses to the number concept of a Weigl-type card sorting test. *J. exp. Psychol.*, 1951, **41**, 23–29.

Heidbreder, E. Towards a dynamic psychology of cognition. *Psychol. Rev.*, 1945, **52**, 1–22.

Hovland, C. I. Computer simulation techniques in behavioral science research. Talk given on *Voice of America*, 1959. (Mimeo.)

Hovland, C. I., & Weiss, W. Transmission of information concerning concepts through positive and negative instances. *J. exp. Psychol.*, 1953, **45**, 175–182.

Lindquist, E. F. *The design of experiments in psychology and education.* Boston: Houghton-Mifflin, 1953.

Smoke, K. L. An objective study of concept formation. *Psychol. Monogr.*, 1932, **42**, (4, Whole No. 191).

(Early publication received October 30, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

REINFORCERS AND REINFORCEMENT: THEIR RELATION TO MAZE PERFORMANCE[1]

WILLIAM BEVAN AND ROBERT ADAMSON[2]

Emory University

Thinking about reinforcement has been marked, for the most part, by a tendency to equate internal process and stimulus agent and a preoccupation with frequency of presentation of the agent as the crucial reinforcement variable. Thus Tolman and Postman (1954), have commented, ". . . it is our view that it is the frequency of S-R-S occurrences and not amounts of reinforcement which primarily affect acquisition" (p. 37). Yet the appearance of more than 25 papers on reinforcer magnitude in the past several years challenges this opinion. A cursory examination of this literature clarifies the need for distinguishing between process and agent. While older studies, like those of Gantt (1936) and Grindley (1929), have indicated a positive connection between performance efficiency and incentive magnitude, the demonstration of contrast effects (e.g. Crespi, 1942; Zeaman, 1949) denies the possibility that the reinforcer-performance relationship is simple and constant. Estes (1956), in referring to the difficulty in generalizing about the nature of this relationship, has concluded that progress toward systematization depends upon the construction of a common scale of measurement. A fruitful approach to this requirement he sees in a ". . . correlation of measures obtained in learning experiments with those obtained in psychophysical experiments" (p. 13). Previous to this, Campbell and Kraeling (1953), studying noxious stimulus reduction as reinforcement, had confirmed the expectation that reduction of shock by a given amount had less effect upon performance with high levels than with low levels of stimulation. They had proposed that some function like Weber's law may be applicable in describing reinforcement.

The concern of the present writers with the magnitude problem derives from an interest in the range of applicability of the approach of frame-of-reference psychophysics. This orientation has provided for the competent prediction of sensory magnitude from an assessment of the intensities of a sequence of stimulus inputs. It has also effectively accounted for anchor and context effects, constancy and contrast, and certain aspects of stimulus equivalence (Helson, 1959). Should it be permissible to assume a reinforcement analogue for the sensory adaptation level, shifts in the effectiveness of reinforcers, contrast effects, the retroaction paradox, adaptation to repeated

[1] The experiments of the present paper were conducted during the summer of 1957 and the fall of 1958 under a grant from the Emory University Research Committee. The writers would like to thank Harry Barker and James M. Burt for their assistance in data collection. Systematization of the data was part of a project carried out during the summer of 1958, while the writers were in attendance at the Interdisciplinary Behavioral Sciences Research Conference, University of New Mexico. This conference was supported by the Behavioral Sciences Division, Air Force Office of Scientific Research, under Contract AF 49 (638)-33.

[2] William Bevan is now at Kansas State University, Manhattan, Kansas; Robert Adamson is at System Development Corporation, Lodi, New Jersey.

reinforcement, the superiority of partial reinforcement, and other phenomena associated with performance efficiency might be brought under a single rubric. Meanwhile, examination of the appropriateness of applying the adaptation-level model in the reinforcement context requires first the demonstration that reinforcement has certain properties. The experiments of the present paper were directed toward identifying several of these prerequisite characteristics.

Experiment I: Scaling of Electric Shock

Adaptation-level theory describes apparent sensory magnitude as a differential between a present stimulus process and an internal norm derived, through a process of pooling, from background and residual (past) stimulation. In this initial experiment, the primary goal was to determine if electric shock, a widely used reinforcer, was scalable in the same fashion as prothetic dimensions to which AL theory already has successfully applied. That shock is scalable is clear. Prior to the present experiment, Swartz (1953) had used shock to the tooth pulp to construct a fractionation scale for pain; and more recently, Stevens, Carton, and Shickman (1958) have described both category and ratio scales for electric shock. The present study, however, is concerned less with delineating the scalar properties of shock, than with demonstrating the generation of an indifference point (AL) dependent upon the relative frequency of the several magnitudes of the stimulus series. This is a requirement of Ex. III.

Procedure. Thirty-seven undergraduate and graduate students each made 100 judgments daily on three consecutive days. For each series of 100 shocks, there were five discrete intensities delivered by a square-wave generator (Applegate stimulator, Model 226B): 1300, 1800, 2300, 2800, and 3300 μamp ac, with a pulse rate of 100 cps. Duration of each presentation, controlled by a Wichita stimulus controller, was 200 msec. The interval between successive presentations was 10 sec. The onset of each shock, manually controlled, was signaled by a small green light, turned on by a Phipps and Bird synchronous clock. Shock electrodes were two 2-cm. silver discs coated with bentonite paste and attached to the left wrist, one on the proximal and one on the distal side.

The psychophysical method was the rating scale version of the absolute method, with nine categories from *very, very weak* through *medium* to *very, very strong*. The usual procedural precautions were observed, *S* being requested to base his judgments only on the shocks presented in the current session. The *S*s were allowed additional categories at either end of the rating scale.

For each of the three sessions, a different order of presentation was used: a rectilinear-control series (A), consisting of each of the five intensities presented 20 times in random order; a positively skewed series (B) with the intensities—weakest to strongest—presented 30, 25, 20, 15, and 10 times, respectively; and another positively skewed series (C) with intensities presented 35, 30, 20, 10, and 5 times, respectively. These orders were presented in the following sequences: ABC, BCA, and CAB, approximately one third of the *S*s being randomly assigned to each sequence. The Indifference Point (median category judgment) assigned to each intensity for each *S* under each condition was computed and averaged through *S*s. The data for each series were fitted by rectilinear, logarithmic, and parabolic functions and the appropriate average for each condition determined. These are presented in Table 1, the right-most column containing the empirically obtained IPs for the several series.

TABLE 1

Averages for Each Stimulus Series, Based on Three Assumed Shapes of Psychophysical Function, and the Corresponding Empirically Obtained Indifference Point

Order	Arithmetic Mean	Geometric Mean	Parabolic Mean	Indifference Point
A	2300	2184	2395	2287
B	2050	1947	2109	2087
C	1900	1817	1947	1900

Results. Inspection of Table 1 indicates, first, that the judgmental responses to shock, like the prothetic dimensions previously subjected to category judgment, are scalable and yielded IPs. Secondly, the IPs for the positively skewed series were less than for the rectilinear series and in the expected order (between-series $F = 9.69$; $df = 2/108$; $P < .001$), indicating the induction of anchor effects. Thirdly, the IPs correspond most closely to the arithmetic means of the stimulus series suggesting the S-R function to be linear. This agrees with Swartz's (1953) report of an approximately linear pain function for stimulation to the tooth pulp. Stevens, Carton, and Shickman (1958) also report that their category data for shock to the fingers could be fit by a straight line but argue, because of the nonlinear relationship between their category and magnitude-estimation data, that this is not indicative of direct proportionality between subjective intensity and shock.

Experiment II: Subjective Intensity and Maze Performance

Extension of the adaptation-level model to reinforcement requires the demonstration that performance efficiency is a function of the apparent rather than the physical intensity of the reinforcing agent. The contrast effects reported in a number of animal experiments with positive reinforcers strongly suggest this to be the case. Both Crespi (1942) and Zeaman (1949) indicate running speed for small amounts of food to drop below an expected level when their *S*s had been previously run to a larger amount, while the opposite held for animals initially run to a small amount and shifted to a larger reward. More recently, Collier and Marx (1959) have reported similar differences in acquisition curves for bar pressing. But the contrast data are ambiguous with regard to the role of subjective intensity, since the criterion trials are, in most studies, run with different physical magnitudes of reinforcer. Further, all have involved changing level of reinforcement during the sequence of acquisition trials, allowing for the possibility that shift may be the significant variable, and thus making plausible explanation in more directly motivational terms. Indeed, Crespi, for example, has described his results in terms of the rodent analogues of elation and depression. Meanwhile, Collier and Marx tested with a common intermediate magnitude. They also adapted with magazine training and tested with bar pressing. The present experiment involves comparing maze learning to a common intensity of shock, after pretest adaptation with different intensities of shock, applied outside the context of the learning situation.

Procedure. Three groups of 17 *S*s were required to judge the intensity of 30 shocks in a pretest situation, using the rating scale of Exp. I. Group A received 30 shocks of 1300 μamp at 10-sec. intervals, Group B, the control group, 30 of 2300 μamp., and Group C, 30 of 3300 μamp. Following the judging series, all *S*s were given a 28-unit, two-choice, bolthead maze problem and required to make five trips around the maze perimeter. Each trip or trial sequence thus involved 28 choices. The maze stylus was manipulated by the right hand, contact with the incorrect bolthead of each pair resulting in the delivery of a single 200-msec. shock of 2300 μamp. through electrodes on the left wrist. The instrumentation for the automatic delivery of reinforcement, as well as for the required programming of shocks of different intensities is described elsewhere (Adamson & Bevan, in press). Choices were spaced at 10-sec. intervals, *S* being paced by the green signal light used in Exp. I. The dependent variable consisted of mean errors for each of the successive trial sequences. These are graphed in Fig. 1 and were subjected to a series of trend analyses (Alexander, 1946) from which the comparisons of interest are summarized in Table 2A.

Results. Inspection of Fig. 1 indicates that maze performance differs

with the apparent intensity of reinforcing shock. On the initial sequence, all three groups performed near a chance level, but with subsequent trips the curves diverge, indicating different rates of acquisition. Acquisition was most rapid in Group A, for which the reinforcement was strong, relative to the magnitude of adapting shock. In contrast, performance was poorest in Group C, for which reinforcement was relatively weak. The over-all trend test performed on all three groups (cf. Table 2A) yielded significant *F*s for both between-group means and between-group slopes sources of variance. Tests performed on the several-paired combinations of the groups indicate Group C to be inferior to both B and A, but B not to differ from A. It is possible that had the difference between adapting and test shock levels been greater for A, its superiority over B might have been statistically as well as graphically demonstrated.

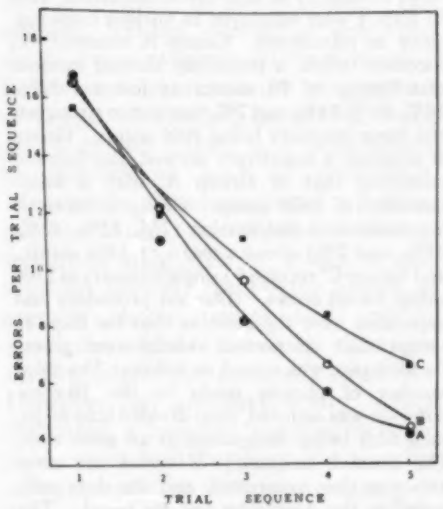


FIG. 1. Performance on a bolthead maze as a function of the relative intensity of reinforcing shock to pretest adaptation shock. For Group A the reinforcing shock was more intense than the pretest shock. For Group B, both were of the same intensity. For Group C the reinforcing shock was weaker. For all groups the reinforcing shock was of the same physical intensity.

TABLE 2

F RATIOS OBTAINED FROM TREND TESTS: EXP. II AND III

Comparison	*df*	*F*	*P*
A. Experiment II			
Over-all Means	2/144	4.96	<.01
Over-all Slopes	2/144	5.81	<.01
C–B Means	1/96	4.05	<.05
C–B Slopes	1/96	7.95	<.01
C–A Means	1/96	15.61	<.01
C–A Slopes	1/96	9.25	<.01
B–A Means	1/96	2.57	>.05
B–A Slopes	1/96	0.001	>.05
B. Experiment III			
Over-all Means	2/135	44.00	<.001
Over-all Slopes	2/135	10.38	<.001
A–B Means	1/100	29.00	<.001
A–B Slopes	1/100	165.00	<.001
A–(C, C′) Means	1/108	86.80	<.001
A–(C, C′) Slopes	1/108	11.80	<.001
B–(C, C′) Means	1/108	8.96	<.001
B–(C, C′) Slopes	1/108	3.66	<.05
C–C′ Means	1/78	.093	>.05
C–C′ Slopes	1/78	.002	>.05

EXPERIMENT III: AVERAGING OF REINFORCEMENT OVER TRIALS

Adaptation-level theory assumes that the internal referent that determines subjective magnitude is the product of pooled residual and background stimulation. After determining the relative influence of residual and background for particular sensory dimensions and psychophysical methods, it is possible to quantitatively specify shifts in the Indifference Point with changes in the relative frequency of presentation of the several series stimuli and with the introduction of anchors (Helson, 1947; Johnson, 1944). It is also possible to predict equivalent patterns of input (Bevan & Darby, 1955). The analogue for the reinforcement concept would

involve predicting the magnitude of change in performance efficiency from changes in the magnitudes comprising a series of reinforcing stimuli as well as identifying a priori equivalent magnitude patterns of reinforcer intensities.

Prerequisite to this is the demonstration that performance efficiency may be related systematically to average reinforcement level. Supportive evidence is found in the report of Logan, Beier, and Ellis (1955) that average running speed did not differ between a group reinforced with five pellets on each trial and a group receiving nine pellets on one half of the trials and one on the alternative half. Also relevant are the findings of Leventhal, Morrell, Morgan, and Perkins (1959) that no preference could be demonstrated for an alley reinforced on every trial over one reinforced with double the amount on half the trials. In the present experiment, the maze performance of a control group receiving a fixed magnitude of shock for every error is compared to that of a group receiving a symmetrical distribution of shocks with the same mean intensity and to groups receiving positively and negatively skewed distributions of shock. It is expected that the first experimental group will not differ significantly from the control, while those receiving greater and less average shock will perform with different efficiency.

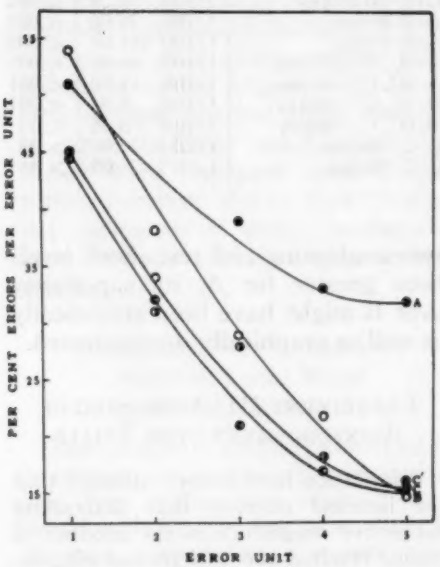


FIG. 2. Performance on a bolthead maze as a function of the mean intensity of reinforcing shock. Group A received the five intensities used in the first experiment, skewed from weak to strong, in the following proportions: 31%, 26%, 22%, 14%, and 7%. Group B received the same intensities, skewed from weak to strong, in these proportions: 7%, 14%, 22%, 26%, 31%. Group C received a symmetrical distribution about the mean intensity in the following proportions: 7%, 22%, 42%, 22%, and 7%. Group C′ received a single intensity (the mean value) on *all* reinforced trials.

Procedure. In order to separate the possible effects of varying intensity over trials from varying frequency of reinforcement, it was necessary to insure that all *Ss* receive the same number of reinforced trials. Accordingly, a more difficult task—a 40-unit, two-choice, bolthead maze problem—was used to insure a specified minimum of reinforced trials. Four groups of *Ss*, 48 in all, were tested. The five shock intensities used in Exp. I were employed in various combinations as reinforcers. Group A received, on incorrect trials, a positively skewed random distribution of 70 shocks as follows: 31%, 26%, 22%, 14%, and 7%, weakest to strongest, the mean intensity being 2000 μamp. Group B received a negatively skewed distribution, mirroring that of Group A, with a mean intensity of 2600 μamp. Group C received a symmetrical distribution (7%, 22%, 42%, 22%, and 7%) about a mean of 2300 μamp., and Group C′ received a single intensity of 2300 μamp. for all errors. The test procedure and apparatus were the same as that for Exp. II except that no pretest shocks were given. Performance was scored as follows: The total number of choices made to the 70-error criterion was counted, then divided into fifths, each fifth being designated as an error unit. The relative frequency of errors per error unit was then computed, and the data subjected to the Alexander test for trend. The results are presented in Fig. 2 and Table 2B.

Results. Examination of Fig. 2 suggests that Group A (lowest mean intensity) gave the poorest performance and Group B (highest mean intensity) the next poorest. The best performance was associated with

the average shock of intermediate intensity. Meanwhile, Groups C and C′, in line with expectation, did not differ significantly from each other. They were thus combined for statistical comparison with Groups A and B. Table 2B indicates each of the three mean reinforcement magnitudes to differ from the remaining two. It would appear, therefore, that performance level relates reliably to average reinforcement level. The establishment of this relationship is necessary to the extension of the pooling model to this problem area. That an averaging effect is tantamount to the pooling of reinforcement over time is, however, still unclear. A definitive demonstration of the latter requires the quantitative prediction of performance associated with some designated average level from the performance levels associated with each of the component intensities that contribute to the average. The present writers currently are conducting a parametric study of runway behavior, with reinforcer magnitudes of 1, 2, 3, or 4 pellets on 25, 50, 75, or 100% of the trials, as a preliminary to attempting such a prediction. Meanwhile, the operation of a pooling mechanism in the present situation would be suggested by a lack of correlation between repetitive errors at specific choice points and specific strengths of shock. Such data are unfortunately not presently available. However, an ancillary study, which indicated a low but significant negative correlation between the number of perseverative errors on maze performance and the magnitude of AL shifts in the judgmental situation, failed to reveal any consistently missed choice points over individuals. Furthermore, four *Ss* of Group C′ achieved errorless performance at the end of the second trial sequence, and thus had only a minimum opportunity for the perseveration of errors. Finally, a sizable proportion of *Ss* stated that they remembered their strongly shocked responses best, while an equally large number reported that the more intense shocks confused them.

Discussion

The extension of a psychophysical model to the area of learning and performance has several distinct values. For example, the pooling model—through assessing the value of the reinforcing agent in terms of its discrepancy from an average—can readily account for the phenomenon of contrast. Similarly, the adaptation effect, associated with repeated presentations of a single-intensity reinforcer, may be explained as follows: If it may be assumed that, prior to the initial presentation of a reinforcer, the magnitude of the internal norm is zero, then, after the first presentation, the norm will assume a value intermediate between zero and the value of the reinforcing agent. With successive presentations of a fixed intensity, the value of the norm will more and more closely approximate the reinforcer, and the discrepancy between the norm and the reinforcing agent will diminish, as will the apparent intensity of the reinforcer. Conversely, extinction trials should have the effect of reducing the norms. If then reinforcing stimuli are reintroduced, their value should be enhanced. Reports that postextinction performance exceeds a pre-extinction asymptote (Adamson, Bevan, & Maier, 1959; Lauer & Carterette, 1957) support this interpretation. The approach applied to the adaptation effect is also readily applicable to the problem of partial reinforcement. Furthermore, the retroaction paradox disappears when reinforcement is thought of as temporally extended.

Summary

The present paper describes three experiments directed toward examining the status

of three assumptions prerequisite to the extension of the pooling model, associated with adaption-level theory, to the area of reinforcement. The assumptions were: (*a*) that *Ss*, presented a common reinforcing agent as a psychophysical stimulus, would behave as they do when faced with the task of scaling a prothetic dimension—intensities would be ordered, PSEs derived, and anchor effects in evidence; (*b*) that maze performance could be related in significant degree to the relative, in contrast to the absolute, intensity of the reinforcer; and (*c*) that maze performance would vary as a function of the average intensity of reinforcer.

The *Ss* were 136 college students given electric shocks of weak-to-mild intensity either as stimuli for absolute judgments or following errors on a bolthead maze. The results of all three experiments confirmed the validity of the assumptions.

References

Adamson, R. E., & Bevan, W. A simple and inexpensive programming device for psychophysical and learning experiments. *J. gen. Psychol.*, in press.

Adamson, R. E., Bevan, W., & Maier, B. Shifts in bar pressing as a function of alternating extinction and reinforcement. *J. gen. Psychol.*, in press.

Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.

Bevan, W., & Darby, C. L. Patterns of experience and the constancy of an indifference point for perceived weight. *Amer. J. Psychol.*, 1955, **68**, 575–584.

Campbell, B. A., & Kraeling, D. Response strength as a function of drive level and amount of drive reduction. *J. exp. Psychol.*, 1953, **45**, 97–101.

Collier, G., & Marx, M. H. Changes in performance as a function of shifts in the magnitude of reinforcement. *J. exp. Psychol.*, 1959, **57**, 305–309.

Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.

Estes, W. K. Learning. *Ann. Rev. Psychol.*, 1956, **7**, 1–38.

Gantt, W. H. An experimental approach to psychiatry. *Amer. J. Psychiat.*, 1936, **92**, 1007–1021.

Grindley, G. C. Experiments on the influence of amount of reward on learning in young chickens. *Brit. J. Psychol.*, 1929, **20**, 173–180.

Helson, H. Adaption-level as frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 1947, **60**, 1–29.

Helson, H. Adaptation-level theory. In S. Koch, (Ed.), *Psychology: A study of a science.* Vol. I. New York: McGraw-Hill, 1959. Pp. 565–621.

Johnson, D. M. Generalization of a scale of values by the averaging of practice effects. *J. exp. Psychol.*, 1944, **34**, 425–436.

Leventhal, A. M., Morrell, R. F., Morgan, E. F., Jr., & Perkins, C. C., Jr. The relation between mean reward and mean reinforcement. *J. exp. Psychol.*, 1959, **57**, 284–287.

Logan, F. A., Beier, E. M., & Ellis, R. A. Effect of varied reinforcement on speed of locomotion. *J. exp. Psychol.*, 1955, **49**, 260–267.

Lauer, D. W., & Carterette, T. S. Changes in response measures over repeated acquisitions and extinctions of a running habit. *J. comp. physiol. Psychol.*, 1957, **50**, 334–338.

Stevens, S. S., Carton, A. S., & Shickman, G. M. A scale of apparent intensity of electric shock. *J. exp. Psychol.*, 1958, **56**, 328–334.

Swartz, P. A new method for scaling pain. *J. exp. Psychol.*, 1953, **45**, 288–293.

Tolman, E. C., & Postman, L. Learning. *Ann. Rev. Psychol.*, 1954, **5**, 27–56.

Zeaman, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

(Received April 30, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

RELATIONSHIP BETWEEN MAS SCORES AND ASSOCIATION VALUES OF NONSENSE SYLLABLES [1]

L. PHILIP TRAPP AND DONALD H. KAUSLER
University of Arkansas

In studies relating anxiety drive level, as measured by the Taylor Manifest Anxiety Scale (MAS), to performance on a verbal learning task, a frequently used material has been nonsense syllables, presented either serially (e.g., Montague, 1953; Saltz & Hoehn, 1957; Sarason, 1956, 1957; Silverman & Blitz, 1956) or as paired associates (e.g., Taylor & Chapman, 1955). A procedure commonly followed in these studies is to select syllables systematically in order to control for task difficulty and/or intratask competition. The specific variables generally manipulated for this purpose are the similarity of the syllables within the list and the association value of these syllables. Similarity is controlled by means of the letter content of the syllables, and association value is usually determined from one of the early calibrations by Glaze (1928) or Hull (1933). The selected syllables then become the learning task for the groups differentiated on the anxiety dimension.

The present study is concerned with association value since it, like MAS score, is an *S* variable. Similarity is of perhaps equal importance as a variable in determining task difficulty and intratask competition, but since it is defined operationally by *E* manipulation, it is essentially a variable of *task structure* independent of *S*. Several studies have regulated task difficulty and/or intratask competition by manipulating the association values of the syllables involved. For example, Montague (1953) employed lists of different association values as one technique for varying intralist interference; Saltz and Hoehn (1957) controlled task difficulty by means of association values while varying intratask competition through manipulation of intralist similarity; and Sarason (1957) altered task difficulty by means of association values in testing the relationship between task difficulty and MAS level. In studies of this nature, it is implicitly assumed that the degree of task difficulty and/or competition is independent of anxiety drive as measured by the MAS. That is, the association values of the syllables are considered comparable for anxious and nonanxious groups at the beginning of the learning trials.

The present study is designed to test the independence of nonsense syllable association with respect to drive level as evaluated by MAS. Since both association to nonsense syllables and score on the MAS are *S* variables, covariance may possibly be found between anxiety level and association value. Davids and Eriksen (1955), employing a word association test, did find a significant positive correlation between MAS scores and productivity of associations, and Kausler and Trapp (1959) stated more explicitly the rationale

[1] This study was supported by a grant from the College of Arts and Sciences Research Fund, University of Arkansas. Dale Wise assisted in collecting and analyzing the data.

and implications for this suspected relationship. If a significant relationship can be demonstrated between MAS scores and association values of nonsense syllables, there would be valid grounds for questioning the methodology in studies which control task properties in terms of association values of the individual items. For example, lists of syllables supposedly equated in difficulty for different levels on the MAS may actually represent lists of varying difficulty for different levels of the MAS.

Method

Subjects. The *S*s consisted of 90 male students, with approximately median ACE scores, in general psychology classes at the University of Arkansas. (Eight *S*s had to be dropped from the study for failure to follow the instructions.) Since some studies suggest that the MAS applied to an unselected population differentiates groups on sex and intelligence, the homogeneity of the present sample eliminated a potential bias. The *S*s were tested during regular class periods.

Association task. A list of 90 nonsense syllables was prepared for group administration in a booklet containing five pages with 18 syllables to the page, arranged in random order. The nonsense syllables were selected on the basis of meeting two criteria. The first requirement was that the association values reflect close agreement between the calibrations made by Hull (1933) and a revision of his work (Trapp & Kausler, 1959). The revision involved a group rather than an individual presentation of syllables. The aim was to choose nonsense syllables with highly stabilized association values, and this seemed best fulfilled by including nonsense syllables that had demonstrated relatively unchanged association values over a 25-yr. period. With very few exceptions, the selected nonsense syllables deviated less than 5% from the values obtained by Hull in spite of the different procedures used in administration. The second requirement was that the list of nonsense syllables contain equal proportions of syllables with low, intermediate, and high association values. The purpose of this requirement was to make possible an analysis of interaction between anxiety level and level of association value. Hence, 30 of the nonsense syllables had association values ranging from 5% to 15% (low values), 30 had values ranging from 20% to 30% (intermediate values), and 30 had values ranging from 35% to 66% (high values).

Procedure. The *S*s were informed that they were participating in a study of the association value of nonsense syllables. They were instructed to look at each syllable briefly, and if the syllable aroused any particular thought or idea, they were to write it down on the separate answer sheet. The use of a template restricted their attention to one syllable at a time.

Approximately two weeks later, *S*s were given the Taylor MAS, disguised as a biographical inventory, by a different *E*. The form used consisted of the 50 anxiety-keyed items and 40 buffer items. The *S*s were informed that *E* was collecting biographical information about students at Arkansas to compare with students elsewhere. Apparently none of the *S*s suspected any relationship between the two parts of the experiment.

Results

For each *S* two sets of measurements were obtained: an MAS score as determined by the Taylor scoring system, and association scores (number of syllables responded to with an association) for the low, intermediate, and high levels of the list, and for the total list. The summary data showing the relationship between MAS score and association scores are given in Table 1. The MAS scores ranged from 3 to 32 with a median of 16. Total association scores for all 90 syllables ranged from 3 to 84 with a median of 41.5 and mean of 43.38. In order to determine the over-all degree of relationship between MAS score and total association score, a correlation ratio was computed between these two variables. The correlation ratio was used since the data suggested a curvilinear relationship. An eta of .41 ($P < .05$) was obtained, which was mainly attributable to the high association scores by *S*s in the upper range on the MAS.

TABLE 1

SUMMARY DATA ON LIST OF NONSENSE SYLLABLES (90 ITEMS) WITH THREE LEVELS OF ASSOCIATION VALUE FOR SCORES ON MAS

MAS Score	*N*	Association Value						All Syllables	
		Low		Intermediate		High			
		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
0–9	11	7.82	8.56	13.09	8.56	18.27	8.99	39.18	24.90
10–11	12	8.00	8.28	12.75	10.39	18.83	9.54	39.58	26.82
12–13	9	8.78	9.87	14.33	11.29	19.89	11.95	43.00	31.94
14–16	10	9.20	8.36	12.40	9.17	18.70	10.57	40.30	27.13
17–18	9	3.67	3.49	8.56	6.72	16.89	9.26	29.11	17.01
19–20	9	15.22	9.68	19.11	11.96	25.22	11.15	59.56	31.05
21–24	10	7.40	9.93	11.50	10.26	17.00	10.61	35.90	29.77
25 & above	12	14.50	10.21	18.83	10.53	25.33	9.77	58.67	29.19
All *Ss*	82	9.40	8.66	13.90	9.17	20.07	8.35	43.38	24.30

Following the practice used in most studies with the MAS, *Ss* were assigned to one of three groups on the basis of their MAS scores. The 31 *Ss* scoring 19 or above constituted the high anxiety group (HA); the 19 *Ss* scoring between 13 and 19 constituted the medium anxiety group (MA); and the 32 *Ss* scoring 13 or below comprised the low anxiety group (LA). The MAS cutting scores were arbitrarily selected, representing the most logical and, considering ties, practical tripartite division of the total distribution of scores. As is apparent from Table 1, the various subgroups did not differ significantly in variability. Differences between the three levels of anxiety on the three levels of association values were then tested by an analysis of variance. The results of the analysis of variance are given in Table 2. Anxiety level was statistically significant ($F = 3.26$, $P < .05$), but the interaction between anxiety level and level of association value was not ($F = .55$). Hence, the effect of anxiety was independent of the level of association value.

Separate t tests were made between means for the three anxiety groups on the total list of syllables. The only significant difference was between Groups HA and MA ($t = 2.32$, $P < .03$, two-tailed test). The t values between Groups HA and LA and between MA and LA were 1.66 and .80, respectively.

The foregoing analyses were made in terms of differences between *Ss* on association scores. A different, and equally important, approach to the problem consists of analyzing the syllables themselves for differences in

TABLE 2

ANALYSIS OF VARIANCE WITH SUBJECT GROUPING FOR DIFFERENCES IN ASSOCIATION VALUES

Source	*df*	*MS*	*F*
Anxiety (A)	2	617.00	3.26*
Ss in same group	79	189.42	
Association value (AV)	2	2353.30	225.41**
A × AV	4	5.79	.55
AV × Pooled *Ss*	158	10.44	
Total	245		

* $P < .05$.
** $P < .01$.

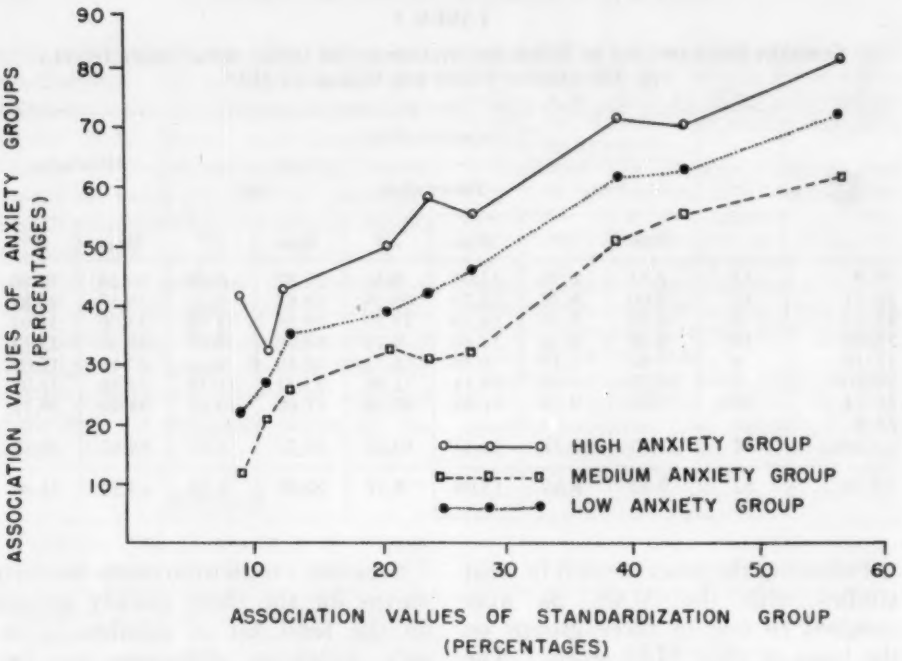


FIG. 1. Association values for groups differentiated on MAS.

association value as a function of anxiety level. For this analysis three association values, calibrated independently for each of the three anxiety groups (HA, MA, and LA) were determined for each of the 90 nonsense syllables. To simplify the analysis, the 90 syllables were grouped in unit blocks of 10 syllables per block and arranged in progressive order of their association values as calibrated from the preliminary standardization study. For each level on the MAS the average association value (percentage of response) was determined for each of the nine blocks of syllables. In Fig. 1 the association values are presented graphically for each MAS group at each point along the continuum representing association value as determined, independently of anxiety level, in the preliminary standardization study.

TABLE 3

ANALYSIS OF VARIANCE WITH SYLLABLE GROUPING FOR DIFFERENCES IN TRANSFORMED ASSOCIATION VALUES

Source	*df*	*MS*	*F*
Anxiety (A)	2	387.63	16.87*
Syllables (blocks of 10) in same group (S)	6	22.98	
Association value (AV)	2	1104.72	150.51*
A × AV	4	2.11	.29
AV × Pooled syllables	12	7.34	
Total	26		

* $P < .01$.

As can be readily observed from an inspection of Fig. 1, anxiety level influences association values at all points on the continuum. The effect is consistently in the same direction with Groups HA, LA, and MA ranking in that order at levels of association value.

In order to test for significant dif-

ferences between groups, the percentages were transformed to arc-sines (following the procedure outlined by Snedecor, 1946) and treated by analysis of variance. The summary of this analysis is presented in Table 3. The F of 16.87 ($P < .01$) permits rejection of the hypothesis that no relationship exists between MAS level and association value of nonsense syllables. Again, no significant interaction effect was found between level of anxiety and level of association value ($F = .29$, $df = 4$ and 12). Separate t tests were again made between anxiety groups. The only significant bifference was detween Groups HA and MA ($t = 2.81$, $P < .02$). The t values between Groups HA and LA and between Groups MA and LA were 1.41 and 1.32, respectively.

Discussion

The data, whether approached from the standpoint of *Ss*' productions or from a direct analysis of the nonsense syllables themselves, indicate that association values of nonsense syllables are a function of anxiety level as evaluated by MAS. Lists of nonsense syllables serially presented, therefore, cannot be assumed to be of equal difficulty at the start of learning trials for groups differentiated on MAS if task difficulty and/or intratask competition is a function of the level of association value. Studies making this assumption have left an important variable uncontrolled, which makes it impossible to infer if the diffential performances obtained between MAS groups reflect differences in drive level, level of task difficulty or intratask competition, or some interaction effect.

From our results, it is apparent that this confounding interaction between MAS level and association value is most pronounced in comparisons between high anxious and moderately anxious *Ss*. Although most studies in this general area ignore the middle anxiety range, an important research goal, nevertheless, is to extend the relationship of anxiety to human verbal learning throughout the entire anxiety continuum. In this extension, an important methodological consideration should be to control for the difference in association value within this range.

The nature of the relationship between association values and MAS poses an interesting theoretical problem and is deserving of some consideration even though it is incidental to the primary purpose of the study. In Fig. 1, it may be recalled, the highest association values were for Group HA, the next highest values for Group LA, and the lowest values for Group MA. The authors found a similar pattern in an independent preliminary study which strongly discounts the possibility of attributing this finding to an artifact of the experimental design. One possible interpretation—accepting for present purposes that MAS is measuring both general drive (Taylor, 1956) and levels of interfering or competing responses (Mandler & Sarason, 1952)—may be advanced as follows: MAS measured drive follows a positively accelerated curve whereas the number or strength of interfering responses project a negatively accelerated curve. Performance, facilitated by drive and inhibited by interfering responses, is conceived as a function of the net differential between the two factors. The largest differential would be among *Ss* in Group HA (drive level having risen sharply but interference level having tended to flatten out) so their performance would be highest. On the other hand, *Ss* in Group MA with only slightly higher drive level than *Ss* in Group LA but with relatively many more interfering responses, would produce the smallest differential and, consequently, reflect lower performance than *Ss* in Group LA.

One other point needs some comment. From Fig. 1 it may be noted that the experimental group as a whole (combining all levels of anxiety) reflected higher association values than the standardization group. This discrepancy can be readily explained in terms of differ-

ences in the task situation for the two groups. The standardization *Ss* were presented with a list of 320 nonsense syllables whereas the experimental *Ss* were presented with only 90 syllables from this list. It would be expected that fewer syllables would elicit a higher percentage of responses.

Summary

The present study investigated the relationship between MAS scores and association values of nonsense syllables. Three groups of male *Ss* (HA, MA, LA) were given 90 nonsense syllables, ranging from low to high association values. A correlation ratio was computed between MAS scores and number of associations, with an obtained eta of .41 ($P < .05$). The association values were grouped in three levels, and differences between the three levels of anxiety on these three levels of association values were tested by an analysis of variance. Anxiety level was statistically significant ($F = 3.26$; $P < .05$) but interaction was insignificant ($F = .55$). Separate t tests were additionally made, with the only significant difference found between Groups HA and MA ($t = 2.32$; $P < .03$). The nonsense syllables were also analyzed for differences in association value as a function of anxiety level. Three association values on blocks of 10 syllables were calibrated independently for each of the three anxiety groups, with significant differences obtained ($F = 16.87$; $P < .01$). A significant t value was found between Groups HA and MA ($t = 2.81$; $P < .02$). Again, no interaction was found. The methodological implications of the general findings were discussed.

References

Davids, A., & Eriksen, C. W. The relation of manifest anxiety to association productivity and intellectual attainment. *J. consult. Psychol.*, 1955, **19**, 219–222.

Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.

Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730–734.

Kausler, D. H., & Trapp, E. P. Methodological considerations in the construct validation of drive-oriented scales. *Psychol. Bull.*, 1959, **56**, 152–157.

Mandler, G., & Sarason, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, **47**, 166–173.

Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.

Saltz, E., & Hoehn, A. J. A test of the Taylor-Spence theory of anxiety. *J. abnorm. soc. Psychol.*, 1957, **54**, 114–117.

Sarason, I. G. Effect of anxiety, motivational instructions, and failure on serial learning. *J. exp. Psychol.*, 1956, **51**, 253–260.

Silverman, R. E., & Blitz, B. Learning and two kinds of anxiety. *J. abnorm. soc. Psychol.*, 1956, **52**, 301–303.

Taylor, J. A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303–320.

Taylor, J. A., & Chapman, J. P. Paired-associate learning as related to anxiety. *Amer. J. Psychol.*, 1955, **68**, 671.

Trapp, E. P., & Kausler, D. H. Revision of Hull's Table of Association Values for 320 selected nonsense syllables. *Amer. J. Psychol.*, 1959, **72**, 423–429.

(Received May 6, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

LATENCY OF INSTRUMENTAL RESPONSES AS A FUNCTION OF COMPATIBILITY WITH THE MEANING OF ELICITING VERBAL SIGNS [1]

ANDREW K. SOLARZ [2]

University of California, Davis

Does the ease of association of an instrumental movement with a verbal stimulus depend upon the compatibility of the movement with the meaning of that verbal stimulus? Such a relation would appear to follow from the two-stage mediational model described by Osgood (1953). He assumes that in the development of meanings of signs, either perceptual or verbal, the sign-stimuli become associated with distinctive portions of the total behavior elicited by significate-stimuli. This portion of the total behavior to things signified is called a *representational mediation process*. It is representational because it is part of the same behavior made to the significate; it is mediational because, through its self-stimulational properties, it can become associated with overt instrumental movements.

Among the classes of responses elicited by both signs and significates are *evaluative reactions*—approaching, avoiding, reaching for, pushing away, and so forth. If the total behavior elicited by a positive or negative significate includes such approach or avoidance movements, then the theory requires that tendencies or dispositions toward such movements must become part of the total mediation process (i.e., part of the meaning of the sign). Furthermore, since the perceptual and vocal signs of positively evaluated objects will often mediate approach movements and those of negatively evaluated objects avoidance movements, one may expect such mediators to tend to elicit overt approach or avoidance movements directly. There are, therefore, two bases in theory to expect easier learning of compatible signs and movements: the evaluative nature of the representational process itself (based on approach and avoidance movements originally made to the things signified) and the generalized association of such mediators with overt instrumental movements.

The purpose of the present study is two-fold: (*a*) to provide a situation in which a two-stage, mediation model is required to account adequately for the data; and (*b*) to provide some evidence for the essentially behavioral nature of mediation processes. A situation is required in which the verbal signs to be used as stimuli have not been previously associated with the particular approach and avoidance movements—otherwise no two-stage interpretation would be necessary. In the present experiment it is assumed that people have seldom, if ever, learned to pull *printed words* toward themselves or to push them away, even though they may have done so in connection with

[1] The study was conducted at the University of Illinois.

[2] The author wishes to thank C. E. Osgood for his many helpful suggestions on the design and preparation of this research. Hans Schmidt deserves special thanks for his collaboration on the statistical analysis of the data.

the things signified. If it can be shown experimentally that it is easier to learn to pull favorably evaluated words (like Fragrant, Kind, and Clean) toward oneself than to push them away, and conversely for unfavorably evaluated words (like Foul, Cruel, and Dirty), then this will be interpreted as evidence for a two-stage mediation model (since it is the meanings of the printed signs, not their physical stimulus properties, that are compatible or incompatible with the overt movements). To differentiate between the two alternative interpretations—compatibility of representational processes or of generalized overt movement tendencies with the new movements in the experimental situation—two response indicators are required: one for the time required to *initiate* the correct movement and the other for the time to *execute* the correct movement. If the differences between compatible and incompatible conditions are found to be restricted to the time to *initiate* the correct movements, then this will be interpreted as evidence that the representational mediation processes themselves include the evaluative characteristics of their behavioral origins.

METHOD

Subjects. The *S*s were 20 male and 20 female undergraduate students from psychology courses at the University of Illinois. They were chosen on a voluntary basis with an inducement of 50 cents for participation.

Apparatus. The apparatus consisted of a movable stage mounted on a runway perpendicular to the ventral axis of *S*. Stimulus cards were inserted behind a vertically sliding door on the stage. The *S* manipulated the door of the stage with a hand lever on the handle base of the stage. As *S* pressed the lever, the door dropped exposing the stimulus card. The stage could be moved 7 in. forward or backward from its center position. White markers on the base of the stage and along the runners allowed for recentralization of the stage.

As the stage door dropped it completed two circuits, starting two .01-sec. Standard Electric timers. Moving the stage ¼ in. off its center position broke one timer circuit and stopped one of the clocks. The second timer continued until the base of the stage made contact with either end of the runway, whereupon the second circuit was broken. In this way two latencies were obtained, one consisting of the time between exposure of the stimulus card and initiation of arm movement, and the other the time for the response movement itself.

Design. Each *S* was told that the experiment was a study of reaction time. A list of 10 words was handed to *S* and he was told, "Look at this list of words and think about the meaning of each word in each pair of words." Three minutes were allowed for this preliminary task. After *S* had been shown how to manipulate the apparatus and told that the listed words would now appear on cards, the following instructions were completed:

> When I say 'ready' I want you to let the door drop, read the word on the card and then move it either toward or away from yourself. After your move I will say either 'wrong' or 'right' according to a prearranged pattern of correct moves for the various words. You are to try to learn which way to move each word as soon as you can and to react with that movement as quickly as you can. The move for each word is always the same.

Five lists were used (see Table 1). Each list contained 10 words, five pairs, which are presumed to be on an evaluative continuum. Each pair contained a positively evaluative word and a negatively evaluative word. The fourth pair in each list was included in

TABLE 1

STIMULUS WORDS USED AS SIGNS

List A	List B	List C	List D	List E
1 sweet	tasty	fragrant	pleasant	fresh
1′ sour	putrid	foul	painful	stale
2 kind	sane	smart	patriot	honest
2′ cruel	insane	stupid	traitor	dishonest
3 valuable	good	nice	clean	pretty
3′ worthless	bad	awful	dirty	ugly
4 near	here	toward	this	close
4′ far	there	away	that	distant
5 happy	gay	joyful	healthy	merry
5′ unhappy	sad	mournful	sick	weary

order to accentuate the compatibility or incompatibility of the sign-movement relationship.

The response required to each word on each list was as follows for the original order of presentation: for the first word, evaluative sign—*pleasant*, the required movement was *toward* the body; for the second word, evaluative sign—*unpleasant*, the required movement was *away* from the body. The required movements for the third and fifth pairs were the same as above; however, the movements required for the second and fourth pairs were opposite, that is, an *away* movement to a *pleasant* word and a *toward* movement to an *unpleasant* word.

Two orders of response presentation were employed. In the reversed order the response movement required to each word was opposite to that in the original order. Consequently, under the original order, six of the sign-movement relation tasks were compatible and four were incompatible; the numbers interchanged under the reversed order. Each *S* under either order of presentation was thus required to learn to make five *toward* and five *away* movements to the words on the given list.

Twenty *S*s served under each of the two orders of response presentation. Four *S*s were used for each list in each order of presentation; they consisted of two men and two women. Each *S* was exposed to the words of only one list throughout the experiment.

A trial consisted of one complete presentation of all 10 stimulus cards. The order of word presentation for any list was varied from trial to trial in accordance with a 12×12 latin square so that each word appeared in a different sequence on each trial. This prevented *S*'s anticipating the next stimulus word.

The reason *S* was not informed of the criterion, two errorless trials, was to prevent his sacrificing speed of response for correctness of response; *S* was informed, however, that the time required for participation would not exceed 1 hr.

Results

Mean latencies were obtained for each *S* for each of the four sign-movement relationships. The latencies for the first trial were not used since the responses here served merely to inform *S* which direction of arm movement was correct for the particular stimulus word. Only latencies of correct responses were used. Finally, these mean values were entered into grand means for the given conditions of the experiment.

TABLE 2

Number of Latency Values Entered into the Final Means Used for Analysis of Variance

Presentation	*N*	Sign-Movement Category[a]			
		++	−+	−−	+−
Original	20	326	203	329	189
Reversed	20	187	267	199	266

[a] First symbol designates sign; second symbol, movement.

Since the errors occurring were not equally distributed over the four categories, the final means are not based upon equal numbers of values, as indicated in Table 2. Also, as noted previously, the number of values entered for compatible and incompatible tasks differed for original and reversed presentation.

If the number of values entered into the final *S* means of each category is a factor in determining the reaction time obtained, a significant interaction involving order of presentation would be expected. If it were the only factor, a reversal in the direction of difference between the final sign-movement means would be anticipated. The presentation of the results shows no such interaction or reversal.

The analysis of variance was applied to the two sets of mean latency scores of individual *S*s for the four sign-movement categories. The results of the analysis are presented in Table 3. As indicated under Clock 1 (time to initiate arm movement) the

TABLE 3

ANALYSIS OF VARIANCE OF SUBJECT MEAN LATENCIES FOR EACH SIGN-MOVEMENT

Source of Variation	*df*	Clock 1 Time to Initiate Arm Movement		Clock 2 Time to Complete Arm Movement	
		MS	*F*	*MS*	*F*
List	4	1176.57		336.80	2.53
Sex	1	191.42		1998.69	15.02***
Order	1	2768.91		333.22	2.50
List × Sex	4	1709.91		454.74	3.42*
List × Order	4	3597.75	1.24	15.86	
Sex × Order	1	64.75		.42	
List × Sex × Order	4	2568.62		76.55	
Ss in same group	20	2895.38		133.04	
Total between *Ss*	39	2490.84		218.69	
Move	1	216.23	1.32	655.69	62.75***
Word	1	556.53	3.40	17.89	1.71
Move × Word	1	3717.17	22.69***	7.61	
Move × Word × Order	1	47.55		4.01	
Word × Move × List	4	915.88	5.59***	14.98	1.43
Word × Move × Sex	1	1328.28	8.11**	18.98	1.82
Word × Move × List × Order	4	427.45	2.61*	5.88	
Other higher interaction effects	47				
Pooled *Ss* × treatments	60	163.81		10.45	
Total within *Ss*	120	176.20		16.99	

* $P = .05$.
** $P = .01$.
*** $P = .001$.

Move × Word interaction is highly significant. Whether initiating an arm movement toward oneself has a shorter or longer latency following stimulus presentation than an arm movement away from oneself is dependent upon whether the sign has a pleasant or unpleasant meaning.

TABLE 4

CLOCK 1 MEANS IN SECONDS FOR MEN AND WOMEN FOR EACH SIGN-MOVEMENT CATEGORY

Move	Men		Women	
	Pleasant Word	Unpleasant Word	Pleasant Word	Unpleasant Word
Toward	.803	.865	.760	.965
Away	.871	.856	.931	.827

The Clock 1 means in seconds for all *Ss* for each sign-movement category were: pleasant-toward = .781; pleasant-away = .901; unpleasant-away = .842; unpleasant-toward = .915. The time needed to initiate a movement to a sign is less when the meaning of the sign and the direction of the movement are compatible than when they are incompatible.

Table 3 also indicates a significant interaction of Word × Move × List for Clock 1. An inspection of the four sign-movement means for each list revealed that List D contributed the greatest total difference, .419 sec.; the shorter latencies being associated with compatible sign-movements. List C showed a reversal in the direction of the difference, a value of .163 sec. Lists A, B, and E

corresponded to List D, showing a range of differences of .199 sec. to .272 sec. in favor of quicker compatible sign-movements.

The Word × Move × Sex interaction is also significant for Clock 1 in Table 3. Table 4 presents the sign-movement means separated for sex. Women contributed significantly more to the direction of the differences than did men.

Order of response presentation itself was not statistically significant nor was it a significant variable in the interaction effects of Move × Word, as indicated in Table 3, Clock 1. However, when the interaction effects of Move × Word are broken down into Lists, order of presentation becomes a significant interacting variable.

Turning attention to the results of the analysis applied to Clock 2 (time of the actual arm movement itself) in Table 3, it may be seen that sex is a significant variable for response time. The mean response time for men was .324 sec., for women .394 sec. The significance of the List × Sex interaction indicates that the speed of response for men and women was differentially affected by the particular list of words employed.

There is no significant relation between sign meaning and direction of movement once the response has been initiated, as indicated by the Move × Word interaction.

The last significant variable in the Clock 2 analysis is the direction of arm movement for all *Ss*. The mean time for arm movement toward the body is .339 sec. and for movement away from the body .379 sec. This is arm movement from the center of the apparatus for a distance of 7 in. toward either the near or far end of the apparatus.

An error-difference score was computed for each *S*. This was obtained by subtracting his mean error on compatible sign-movement from that on incompatible sign-movement. For the group, the mean error-difference score was .468 ($t = 3.72$, $df = 39$, $P < .01$). Separating for sex, the mean error-difference scores were, men .629 ($t = 3.14$, $df = 19$, $P < .01$), women .307 ($t = 2.07$, $df = 19$, $P > .05$). The direction of the differences was in favor of fewer errors on compatible tasks.

The final analysis consisted of a comparison of the combined error scores for men (mean error score on compatible tasks *plus* that on incompatible tasks) with the combined error scores for women. The difference between the means (men 1.222, and women .829) was not significant ($t = 1.81$, $P > .05$).

Discussion

Recent studies (Fitts & Deininger, 1954; Fitts & Seeger, 1953) have indicated shorter latencies and fewer errors when the stimulus of a given stimulus set was spatially compatible with the required response. Stimuli consisted of spatial arrangements of lights or symbols. The results of the present study point out that compatibility extends beyond objective S-R spatial congruity to the subjective level of the compatibility of the sign-meaning with the response-meaning.

In terms of Osgood's (1953) two-stage mediational model of behavior, the learning of movements incompatible to signs would involve some degree of change, either temporary or permanent, in the representational mediation process associated with the stimulus; that is, the sign would assume a postexperimental shift in meaning, particularly within the experimental setting. When the signs and movements are compatible, learning

would consist of a strengthening of the already existing association between sign and mediation process, and a strengthening of the required response within the already associated hierarchy of responses. In this connection note the finding of Solley, Jackson, and Messick (1957) that *S*s, when asked to anticipate, made more guesses for cards containing profiles which had previously been moved "toward" them than for the cards containing profiles without this prior "toward" association. Their finding, as the present one, suggests a perceptual relationship which may serve as reinforcement particularly as regards mediation processes.

The finding of shorter latencies and fewer errors for compatible sign-movements suggests that *S*s came into the experiment with established mediation processes which could be evoked by the proper signs. The existing mediation process associated with a particular evaluative sign was evoked by that sign regardless of the compatibility or incompatibility of the required movement. Osgood states, "Stimulus patterns associated with the same stimulus objects will acquire common mediating reactions and hence will become signs having similar significance" (Osgood, 1953, p. 402). Any particular mediation process is associated with a hierarchy of overt responses, the selection of which is partially determined by the immediate demands of the environment. It is assumed that the hierarchy available to a mediation process, evoked by pleasant or unpleasant signs, contains either responses serving to decrease the distance (sign pleasant) or increase the distance (sign unpleasant) between oneself and the given sign.

Attention is called to the fact that performing incompatible sign-movements did not involve the time to execute the instrumental response (arm movement) itself. One would expect on the basis of animal experiments (Brown, 1948; Bugelski & Miller, 1938) that the rate and strength of an overt response would be affected when an approach or avoidance movement is made to an incompatible evaluative sign. Hovland and Sears (1938) found "blocking and compromise responses" predominating as modes of resolution to approach-avoidance and avoidance-avoidance conflicts. If overt movements contained the evaluative characteristics due to past associations, then the generalized overt movement tendencies to incompatible signs should have resulted in longer times of execution. This was not the case. The results support Osgood's view that the representational mediation processes include the evaluative characteristics of their behavioral origins.

Summary

Twenty men and 20 women served in a study of the relation between arm movements, toward or away from the body, and positively or negatively evaluative signs (words like Smart, Stupid, Happy, etc.). Printed word cards were presented on a movable stage and *S* was required to learn to move the signs according to a prearranged assignment of signs to compatible or incompatible movements. Five lists of 10 words each were employed in the experimental design.

Results revealed shorter latencies of initiating a "toward" or "away" movement to a compatible sign than to an incompatible sign. The compatibility variable did not affect the time of execution of the arm movement itself. In addition, *S*s committed more errors in learning incompatible tasks than compatible ones. Differences between men and women in performance were found on latency measures and on errors of response. The results were interpreted in terms of Osgood's theory of "meaning as a representational mediation process."

References

Brown, J. S. Gradients of approach and avoidance responses and their relation to motivation. *J. comp. physiol. Psychol.*, 1948, **41**, 450–465.

Bugelski, R., & Miller, N. E. A spatial gradient in the strength of avoidance responses. *J. exp. Psychol.*, 1938, **23**, 494–505.

FITTS, P. M., & DEININGER, R. L. S-R compatability: Correspondence among paired elements within stimulus and response codes. *J. exp. Psychol.*, 1954, **48**, 483–492.

FITTS, P. M., & SEEGER, C. M. S-R compatibility: Spatial characteristics of stimulus and response codes. *J. exp. Psychol.*, 1953, **46**, 199–210.

HOVLAND, C. I., & SEARS, R. R. Experiments on motor conflict and their modes of resolution. *J. exp. Psychol.*, 1938, **23**, 477–493.

OSGOOD, C. E. *Method and theory in experimental psychology*, New York: Oxford Univer. Press, 1953.

SOLLEY, C. M., JACKSON, D. N., & MESSICK, S. J. Guessing behavior and autism. *J. abnorm. soc. Psychol.*, 1957, **54**, 32–36.

(Received May 11, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

EFFECT OF DISCRIMINATION TRAINING ON AUDITORY GENERALIZATION

HERBERT M. JENKINS

Bell Telephone Laboratories, Incorporated

AND ROBERT H. HARRISON

Massachusetts Mental Health Center

Two types of training procedures have been used in studies of stimulus generalization. In the nondifferential procedure a response is simply reinforced repeatedly in the presence of, or closely following the presentation of, a stimulus of a particular value. In the differential procedure, on the other hand, a discrimination is developed between two stimulus values by reinforcing the response to one value and extinguishing the response when it occurs to the other value.

Lashley and Wade (1946) argued that differential training was required for the development of a gradient of generalization. The several experiments which faced this issue (Grandine & Harlow, 1948; Grant & Schiller, 1953; Grice, 1948) have demonstrated, however, that a sloping gradient of generalization can be obtained following a nondifferential training procedure.

In experiments comparing gradients of generalization following training with these two procedures (Hovland, 1937a; Schlosberg & Solomon, 1943); it was found that differential training produced a steeper gradient. Although the differential training involved nonreinforcement with respect to a stimulus that differed from the reinforced stimulus along the dimension on which the gradient was obtained, it is not clear that the effect depends entirely on this feature. The purpose of the present experiments was to examine the possibility that differential training not involving an explicit discrimination of this kind also produces a steeper gradient.

The method was similar to the one used by Guttman and Kalish (1956). The key-response of the pigeon was reinforced in the presence of a pure tone. Generalization gradients along the dimension of frequency were compared following differential or nondifferential training. Differential training was used to produce a discrimination between the presence of tone and its absence; it did not involve explicit training to discriminate frequencies. The expectation that training of this form sharpens the gradient of generalization was based on the following reasoning. The establishment of a discrimination between presence and absence ensures that the tone has discriminative control over the response. This provides the potential for an increase in discriminative control by any of the several properties of the tonal

[1] This research was done at the Massachusetts Institute of Technology and was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under Contract AF 18 (603)-85. The present article is a condensed version of the final report on that contract; TN 58-443, ASTIA Document 158248. Reproduction for any purpose of the United States Government is permitted.

The authors wish to express their appreciation to Carl Hovland for his helpful criticisms of a draft of this article.

stimulus; properties which might be expected to include frequency as well as intensity, rate of interruption, and so on.

EXPERIMENT I

Method

Subjects. The *S*s were eight male, white Carneau pigeons maintained by restricted feeding at 80% of their body weight under free feeding.

Apparatus. A Skinner automatic key-pecking apparatus was used. The general features of the apparatus have been described elsewhere (Ferster & Skinner, 1957). The reinforcement was a 4-sec. period of access to a tray of mixed grain. The occurrence of a reinforcement was signaled by turning off the illumination of the response-key, and by turning on illumination for the food tray.

The auditory stimuli were pure tones sounded by a 5-in. speaker suspended 3 in. behind the response-key. The tone was interrupted for 0.25 sec. in each second in order to reduce possible adaptation to a steady tone. The amplification of tones at different frequencies was adjusted to give a sound pressure level of 70 decibels re 0.0002 dynes/cm.^{2} as measured with the aid of a calibrated microphone.

Preliminary training. Four sessions of preliminary training were given. The key-response was conditioned using a method of successive approximations and then placed on a variable interval (VI) schedule of reinforcement. The tone which served as the reinforced stimulus, or S^D, was presented as soon as several responses occurred. The third and fourth sessions were divided into 25 working periods of 33-sec. duration by the introduction of 7-sec. blackout intervals. In each working period, the S^D was presented and responses were reinforced on the VI schedule which delivered from 1 to 3 reinforcements with an average of 1.7 reinforcements.

Procedure. In the nondifferential training method, sessions consisted of 25 working periods in which a tone at 1000 cps (S^D) was presented and responses were reinforced on the VI schedule. Three *S*s received 10, 20, or 40 daily sessions of nondifferential training.

In the differential training sessions, the working periods were of two types: S^D periods as in nondifferential training, and S^Δ periods in which no tone was presented and responses were nonreinforced. Training sessions consisted of 25 S^D periods, and a number of S^Δ periods which was increased from 25 in the first session to 125 in later sessions. The S^D and S^Δ periods were presented in a random order. Daily training sessions were continued until the average rate of response in the S^D periods was at least four times greater than in S^Δ periods in each of two successive sessions. When testing the discrimination against this criterion, the sessions consisted of 25 S^D periods and 50 S^Δ periods. Five *S*s received differential training. From six to nine sessions were required to reach the criterion of discrimination.

In testing for generalization, the stimuli were presented for 33-sec. periods separated by 7-sec. blackouts as in training. Eight test stimuli were used. They consisted of seven tones approximately equally spaced along a logarithmic scale of frequency. Their frequencies were 300 cps, 450 cps, 670 cps, 1500 cps, 2250 cps, 3500 cps, and 1000 cps (S^D in training). The eighth stimulus was no tone. In a test session, the eight stimuli were presented in eight blocks, making a total of 64 presentations. The order of presentation was governed by 8×8 latin squares so that each stimulus appeared once in every ordinal position in the block.

Tests for generalization were conducted without reinforcement except where noted. Three tests for generalization separated by single retraining sessions were usually made. The procedure during retraining was the same as that used during training.

Results

Nondifferential training. The percentage of responses made to each test stimulus following nondifferential training is shown in Fig. 1. Prior to the generalization tests, S_{70} received 10 training sessions, S_{71} received 20,

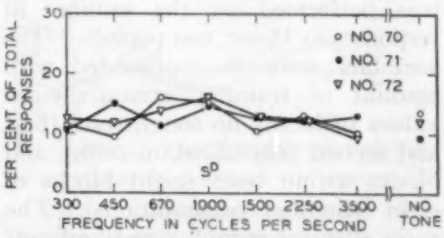


FIG. 1. Generalization gradients following nondifferential training with a 1000-cps tone as S^D. Individual gradients are based on the means of several generalization tests (see text).

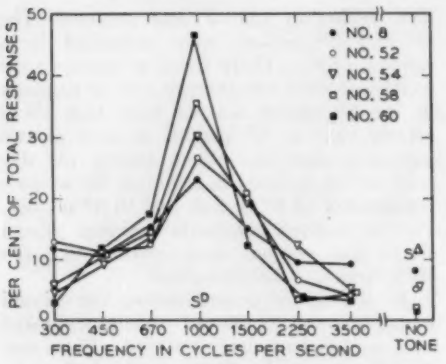


FIG. 2. Generalization gradients following differential training with a 1000-cps tone as S^D; no tone as S^Δ. Individual gradients are based on the means of three generalization tests.

and S_{72} received 40. Gradients are based on three tests on S_{70} and S_{71} and on two tests for S_{72}. Tones widely separated in frequency from the S^D, as well as no tone, were responded to in approximately equal measure. Increasing the amount of nondifferential training from 10 to 40 sessions did not have a marked effect upon the gradient.

There was some evidence for a very shallow gradient of response strength. An average of the individual curves would show the highest percentage of response at the S^D with a slight but systematic decrease at tones of higher or lower frequency.

A four-way analysis of variance was performed on the number of responses in 33-sec. test periods. The variables were *S*s confounded with amount of training, stimuli (eight values including no tone), tests (first and second generalization tests), and blocks within tests (eight blocks of eight stimulus presentations). The main effect of stimuli was significant at the .05 level ($F = 2.38$; 7/273 *df*). Since the simple interactions involving stimuli were not significant, the small main effect of stimuli provides some evidence for a shallow gradient following nondifferential training. Large and highly significant main effects were found for tests ($F = 229.54$; 1/273 *df*, $P < .001$) and for blocks within tests ($F = 36.36$; 7/273 *df*, $P < .001$). The effect due to tests reflects a sharp decline in the number of responses from the first to the second test. This was expected on the basis of the well established finding that extinction occurs more rapidly in successive extinctions following reconditioning (Bullock & Smith, 1953). The effect of blocks reflects a decline in the rate of response due to nonreinforcement during the tests of generalization.

Differential training. Generalization gradients following differential training are shown in Fig. 2. During the test sessions, S_8, S_{52}, and S_{54} received four additional presentations of the S^D in which one reinforcement was delivered. These presentations were excluded from the data. Differential training resulted in well defined gradients along the dimension of frequency. The percentage of responses to extreme frequencies (300 cps or 3500 cps) was only slightly greater than to no tone. The possibility that tones at these frequencies were inaudible, can be ruled out on the basis of data to be presented from Exp. II.

The analysis of variance for the differential group was based on S_{58} and S_{60} which received no reinforcement during the tests for generalization. Three tests for generalization were included in this and subsequent analyses. As is apparent from Fig. 2, the effect of stimuli was large ($F = 47.74$; 7/273 *df*, $P < .001$). The effects of blocks within tests ($F = 8.68$; 7/273 *df*, $P < .001$) and of tests ($F = 14.05$; 2/273 *df*, $P < .001$) again reflect, respectively, extinction within tests,

and more rapid extinction in successive tests. Significant interactions were obtained for the effects of stimuli with blocks ($F = 1.80$; 49/273 *df*, $P < .01$), and of stimuli with tests ($F = 2.28$; 14/273 *df*, $P < .01$). The interaction effects were due in part to a progressive sharpening of the generalization gradient both from the beginning to the end of a test, and in successive tests. Similar trends are examined in more detail with the results of Exp. II.

Experiment II

Experiment II was designed to extend the comparison of generalization gradients obtained following differential or nondifferential training to a case in which responses are reinforced to tones at two frequencies (cf. Kalish & Guttman, 1957).

Method

The same method was used as in Exp. I except for the details noted below.

Procedure. Nondifferential training sessions consisted of 30 S^D periods. A 450-cps tone was presented in 15 of these periods; a 2500-cps tone was presented in the remaining 15 periods. The tones were presented in a random order. Three *S*s received 10, 20, or 40 daily sessions of nondifferential training.

The differential training procedure involved the same S^D periods, and in addition, S^Δ periods in which no tone was presented and responses were nonreinforced. The number of S^Δ periods was increased from 30 in the early training sessions to 150 in later sessions. This procedure does not involve explicit training to discriminate between the two tones used as an S^D.

Daily sessions were continued until the criterion of discrimination was reached (4 to 1 ratio of S^D to S^Δ rate of response) for each tone with respect to no tone. Six *S*s received differential training. From 7 to 11 training sessions were required to reach the criterion of discrimination.

The tests for generalization were made with the following stimuli: tones at frequencies of 250 cps, 450 cps (S^D), 670 cps, 1000 cps, 1500 cps, 2500 cps (S^D), 4000 cps, and no tone. Generalization tests were conducted without reinforcement. Three tests were made with a single retraining session between tests.

Results

Nondifferential training. The gradients obtained following nondifferential training are shown in Fig. 3. Prior to the generalization tests, S_{29} received 10 training sessions, S_{30} received 20, and S_{36} received 40. The gradients are similar to those obtained following nondifferential training with respect to a single frequency in that they show a nearly equal percentage of response to all test stimuli, including no tone. Again, it is apparent that increasing the amount of training did not greatly alter the result. As in Exp. I, there was some evidence for a shallow gradient; each *S* had a local peak in response strength at 450 cps, although the curves over all test stimuli were irregular.

An analysis of variance for this condition shows effects similar to those for the nondifferential condition in Exp. I. The effect of stimuli was again significant ($F = 3.23$; 7/448 *df*, $P < .01$), and the simple interactions of stimuli with blocks or tests were not significant.

Differential training. Generalization gradients following differential training are shown in Fig. 4. Differ-

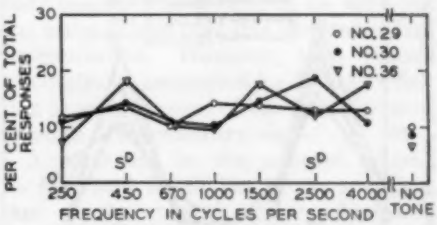


FIG. 3. Generalization gradients following nondifferential training with a 450-cps tone and a 2500-cps tone as S^Ds. Individual gradients are based on the means of three generalization tests.

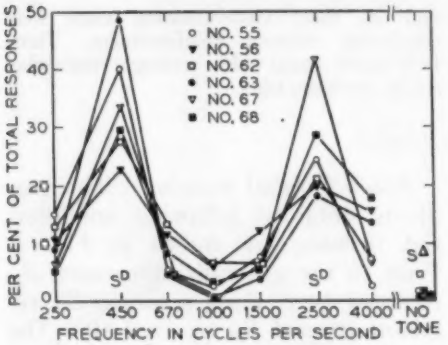


FIG. 4. Generalization gradients following differential training with a 450-cps tone and a 2500-cps tone as S^Ds; no tone as S^Δ. Individual gradients are based on the means of three generalization tests.

ential training resulted in much sharper gradients with peaks in response strength occurring at the frequencies reinforced in training. The effect of differential training in sharpening the gradient is consistent with the effect obtained in Exp. I.

The gradients with peaks near the extremes of the frequency range permit us to reject an interpretation of the low response strength obtained in Exp. I at the extreme frequencies of 300 cps and 3500 cps (see Fig. 2) as resulting from the inaudibility of these frequencies to the pigeon. It is also of interest to note that, in the present experiment, response strength at 1000 cps was only slightly greater than that for no tone; a stimulus which approximated silence. Although these observations do not eliminate the possibility that differences in loudness affected the gradients obtained, they make it appear unlikely that such differences were a major source of variance.

An analysis of variance shows effects that parallel those obtained with differential reinforcement in Exp. I with some exceptions. In particular, the interaction of *S*s with stimuli was highly significant ($F = 5.44$; 35/973 *df*, $P < .001$), whereas it was not significant in Exp. I. The major source of this interaction was in the difference among the *S*s in the S^D (450 cps or 2500 cps) at which they responded most frequently in the generalization tests. A change in the distribution of responses to the test stimuli both in successive blocks within tests and in successive tests was reflected in significant interactions of stimuli with blocks ($F = 2.07$; 49/973 *df*, $P < .01$) and of stimuli with tests ($F = 4.01$; 14/973 *df*, $P < .001$). These effects were also obtained in the differential group in Exp. I, and in both cases, represent

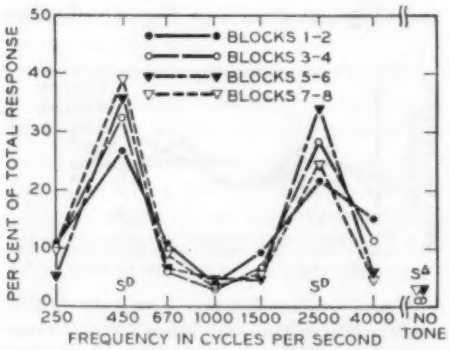


FIG. 5. Average generalization gradients for successive tests; Exp. II.

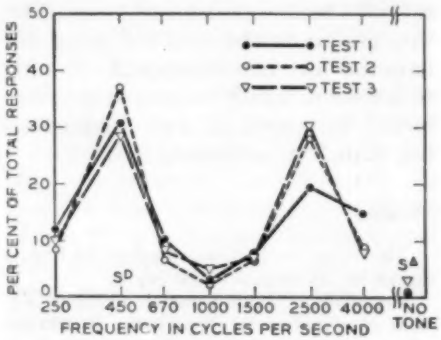


FIG. 6. Average generalization gradients for successive blocks within tests; Exp. II.

a general sharpening of the gradient from the first to later tests, and within the course of a single test. Figure 5 shows the mean percentage of responses to the test stimuli in successive tests. Figure 6 plots these data for successive blocks within the tests.

Discussion

The experiments provide clear evidence that differential training can greatly sharpen a gradient of generalization even when this training does not involve an explicit discrimination between stimulus values along the dimension of the gradient. The conclusion stated by Lashley and Wade (1946, p. 74): "the 'dimensions' of a stimulus series are determined by a comparison of two or more stimuli and do not exist for the organism until established by differential training" is weakened by this result if the phrase "two or more stimuli" is taken to mean stimuli having values on the dimension of the gradient. It now seems likely that the absence or near absence of a sloping gradient along some dimension of a stimulus does not imply that training to discriminate values on this dimension is required. The present results suggest that the failure to obtain such a gradient may be due to a lack of control over the occurrence of the response by the experimental stimulus. When presenting or removing the stimulus has little or no effect upon the occurrence of the response, we can only conclude that the stimulus, and consequently any dimension of the stimulus, is irrelevant to the response.

The results pose a theoretical question: how account for the effectiveness of nonreinforcement in the absence of the tone in placing the tone in control of the response? A discussion of stimulus generalization by Hull (1952, p. 59 ff) is relevant here. He pointed out that in addition to the experimental stimulus, incidental stimuli (e.g., stimuli arising from constant features of the apparatus) accompany the reinforcement of the response. When training is nondifferential, these stimuli are repeatedly reinforced and may predominate in control of the response with the result that the gradient of generalization observed upon varying the experimental stimulus is flat or nearly so. When, on the other hand, training is differential, these incidental stimuli are both reinforced in the presence of S^D and nonreinforced in its absence. This reduces their effectiveness in controlling the response. Hull assumed, as does Restle (1955), that stimuli accompanying reinforced and nonreinforced responses adapt out, or in some manner become nonfunctional. The resulting increase in the relative weight of the experimental stimulus in controlling the response makes it possible to obtain a sharply sloping gradient of generalization along some dimension of the experimental stimulus. It is of interest to note that, according to this account, differential training involving neighboring stimulus values would also weaken the control by incidental stimuli. The effect of that procedure in sharpening the gradient might be largely due to this factor rather than due to the presumed development of a counter gradient of conditioned inhibition centered around the value of the nonreinforced stimulus. In an experiment by Reinhold and Perkins (1955) it was found that differential training on a feature of a compound stimulus unrelated to the feature used in testing for generalization produced a gradient with a steeper slope than did nondifferential training. They suggested that a set to discriminate was developed by differential training and that this resulted in less generalization. However, their results could also be accounted for by the reduction in effectiveness of incidental stimuli through differential training.

A weakness in the account offered by Hull and Restle lies in the specification of the conditions for adaptation of incidental stimuli. It is not sufficient to state that a stimulus accompanying both the reinforcement and nonreinforcement of a response adapts out, since in the present experiment responses to the S^D were reinforced only inter-

mittently, yet with differential training the S^D showed rather precise discriminative control over the response.

Regardless of the interpretation placed upon the present results, it is clear that a property of a stimulus that shows little or no influence on a response after prolonged nondifferential training may gain rather precise discriminative control by means of differential training to discriminate the presence from the absence of the stimulus.

The auditory generalization gradients may be compared with the visual generalization gradients along the dimension of wavelength obtained by Guttman and Kalish (1956) using pigeons. In contrast with the present results, these investigators found steep, well-defined gradients following nondifferential training. Since the training procedures appear to be the same in all important respects it may be concluded that the difference lies in the use of a visual—as compared with an auditory stimulus. It does not appear that a lack of physiological capacity to detect or to make discriminations among auditory stimuli can account for the contrast in the results obtained with visual and with auditory stimuli. The more important role of visual stimuli in the normal activities of pigeons may be a factor. In locating food or in flying, visual discriminations are more heavily involved.

It should be noted, however, that the visual stimulus appeared directly on the response key, whereas the auditory stimulus was diffuse and probably unlocalized. Nondifferential training with a visual stimulus presented in an analogous manner, as an over-all illuminant, might also result in a flat generalization gradient.

Auditory generalization gradients following differential training and visual generalization gradients following nondifferential training are similar in some respects. In each case the typical gradient is negatively accelerated and approximately symmetrical. However, the visual gradients were plotted on a linear scale of wavelength, while the auditory gradients were plotted on a logarithmic scale of frequency. The approximate symmetry of the auditory gradients when plotted in this way may be related to the fact that human discriminability of frequency changes is proportional to the logarithm of frequency for tones above 1000 cps.

The present results, showing that the generalization gradients became sharper in the course of testing, are at variance with those reported by Kalish and Guttman (1957), who found no systematic change in the relative frequencies of response to test stimuli in the course of extinction. Some investigators have found sharpening (Brown, 1942; Hovland, 1937b; Littman, 1949), as in the present case, while others have found that gradients flatten during extinction (Grant & Schiller, 1953; Reinhold & Perkins, 1955). The question of what determines how extinction affects a gradient of generalization is still open.

Summary

An operant conditioning method was used to obtain auditory generalization gradients along the frequency dimension of pure tones for the pigeon. Two procedures for training prior to testing for generalization in extinction were compared. In the differential procedure responses were reinforced in the presence of a tone and nonreinforced in the absence of a tone so that an explicit discrimination between presence and absence was formed. In the nondifferential procedure responses were simply reinforced in the presence of a tone.

In Exp. I the S^D was a tone at 1000 cps. In Exp. II equal reinforcement was given with respect to two S^Ds: a tone at 450 cps and one at 2500 cps. Test stimuli were tones higher and lower in frequency than the S^D (or S^Ds) and the absence of a tone.

Gradients of generalization following nondifferential training were nearly flat; that is, tones of different frequency, as well as the absence of a tone, were responded to about equally. On the other hand, well-defined gradients of generalization with relatively steep slopes were obtained following differential training. When responses were reinforced in the presence of two tones of different frequency, the gradients showed two peaks in the strength of response, one at each of the frequencies reinforced in training. To a

first approximation, the gradients were symmetrical about the S^D or S^Ds when plotted on a logarithmic scale of frequency.

It is concluded that training to discriminate the presence from the absence of a stimulus can greatly enhance the amount of control exerted over a response by a specific property of the reinforced stimulus.

References

Brown, J. S. Generalized approach and avoidance responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.

Bullock, D. H., & Smith, W. C. An effect of repeated conditioning-extinction upon operant strength. *J. exp. Psychol.*, 1953, **46**, 349–352.

Ferster, C. B., & Skinner, B. F. *Schedules of reinforcement.* New York: Appleton-Century-Crofts, 1957.

Grandine, L., & Harlow, H. F. Generalization of the characteristics of a single learned stimulus by monkeys. *J. comp. physiol. Psychol.*, 1948, **41**, 327–338.

Grant, D. A., & Schiller, J. J. Generalization of the conditioned galvanic skin responses to visual stimuli. *J. exp. Psychol.*, 1953, **46**, 309–313.

Grice, G. R. The acquisition of a visual discrimination habit following responses to a single stimulus. *J. exp. Psychol.*, 1948, **38**, 633–642.

Guttman, N., & Kalish, H. I. Discriminability and stimulus generalization. *J. exp. Psychol.*, 1956, **51**, 79–88.

Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148. (a)

Hovland, C. I. The generalization of conditioned response. IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 161–176. (b)

Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

Kalish, H. I., & Guttman, N. Stimulus generalization after equal training on two stimuli. *J. exp. Psychol.*, 1957, **53**, 139–144.

Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72–87.

Littman, R. A. Conditioned generalization of the galvanic skin reaction to tones. *J. exp. Psychol.*, 1949, **39**, 868–882.

Reinhold, D. B., & Perkins, C. C., Jr. Stimulus generalization following different methods of training. *J. exp. Psychol.*, 1955, **49**, 423–427.

Restle, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62**, 11–19.

Schlosberg, H., & Solomon, R. L. Latency of response in a choice discrimination. *J. exp. Psychol.*, 1943, **33**, 22–39.

(Received May 11, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

EFFECTS OF EXTINCTION TRIALS ON DISCRIMINATION REVERSAL [1]

M. R. D'AMATO AND H. JAGODA

New York University

In recent years, considerable interest has been shown in discrimination reversal learning. One important feature of such learning is that, in general, reversal learning (i.e., the first reversal) takes longer than the learning of the original discrimination, at least in rats (Dufort, Guttman, & Kimble, 1954; Kendler & Lachman, 1958). The present research is addressed to this problem. Previous observations, and certain theoretical considerations, led us to suspect that reversal learning was, in part, retarded by the "cue" function of the reward present during the reversal trials. Quite often during reversal learning, animals were observed to revert to the original discrimination habit after receiving a reinforcement on a reversal choice. This behavior seemed theoretically reasonable from the standpoint of the r_g mechanism, or even from some notion as the "aftereffects of reward" as a stimulus factor. It appeared plausible that extinction of the original discrimination was impeded by this cue function of reward, and that, conversely, if reward were removed from the situation, at least for a time, reversal learning might be facilitated.

To test this hypothesis we ran two groups of *S*s on a brightness discrimination problem. Group R was reversed in the normal way upon reaching criterion on the original discrimination. The experimental group, Group ER, first underwent a series of extinction trials after meeting the discrimination criterion. That is to say, Group ER learned the discrimination, underwent 60 extinction trials, and then entered reversal learning. Actually, the extinction trials may be considered reversal learning trials without reward. Our hypothesis then was that the extinction trials would, owing to the elimination of reward from the situation, lead to a more rapid extinction of the original discrimination, and that upon reinstitution of reward reversal learning would therefore occur faster in Group ER than in Group R, for which all reversal trials were rewarded.

METHOD

Subjects. The *S*s were 30 experimentally naive male albino rats, 60–90 days of age.

Apparatus. The apparatus consisted of three units of an automatic learning-discrimination apparatus (ALDA) recently developed at New York University.[2] Essentially, ALDA consists of a symmetrical Y maze (Fig. 1), with remote programming, switching, and recording apparatus. The presently available discriminanda are restricted to variable intensities of flashing (90 flashes/min.) and steady illumination in the arms of the Y maze.

A trial proceeds as follows. Assume that a brightness discrimination problem has been

[1] This research was facilitated by Grant M-2051 from the National Institute of Mental Health.

[2] The senior author is responsible for the design of the apparatus described and the supervision of its construction, financial support arising from a grant from New York University. J. Wilson and R. Lachman contributed imaginatively, and thanks are due R. Berryman for originally suggesting the symmetrical Y maze. H. H. Kendler provided essential support and encouragement throughout the project.

A more complete description of the apparatus is available from the senior author upon request.

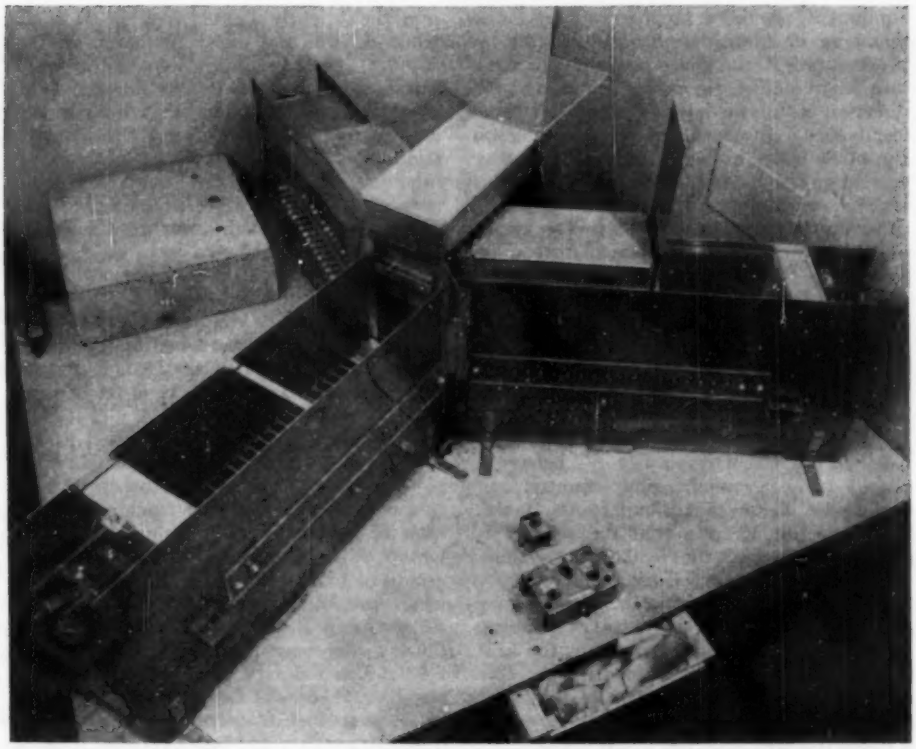


FIG. 1. The symmetrical Y maze of ALDA. The internal dimensions of each arm are: 17 in. long; 5 in. wide; 5.25 in. high. The discriminanda are presented through the translucent plastic of the lid and rear wall.

programmed, with S_+ being the brightly lighted arm and S_- the opposing dimly lighted arm. The *S* is placed into one of the arms, say Arm 1, and the starting button is pressed. The flashing light ("starting" signal) comes on in Arm 1, and at the same time S_+ occurs in one of the other arms, say Arm 2, and S_- appears in Arm 3. At the end of the starting interval (I_4), the doors open and *S* is confronted with the two differently illuminated arms.

Assume *S* makes a correct choice, i.e., he enters the brightly lighted arm. As soon as he penetrates the arm beyond a certain point, a low amperage transistorized contact relay activates, resulting in the closing of the doors and the stopping of the running-time timing mechanism. (Running time is measured from the opening of the doors until their closing.) However, nothing further happens until *S* starts taking his reward (water, sucrose solution, etc.). The first lap of the reward starts the reward interval (I_1) timing, at the end of which the reward dipper retracts out of sight. As soon as I_1 is finished the "stimulus interval" (I_2) starts timing. The purpose of I_2 is to provide a controllable period during which the discriminanda are present but reward is absent. I_3, the "intertrial interval," follows upon I_2. During the intertrial interval the illumination in the maze may be changed in accordance with one's requirements. For example, in the present study all lights were extinguished during the intertrial interval. At the end of I_3, I_4 (the starting interval) begins, indicating the start of a new trial. Each trial runs off as described in accordance with the program, and when the last trial is finished the apparatus "homes out," plunging *S* in darkness. If *S* makes an incorrect choice, or if the trial is programmed to be nonreinforced, I_1 starts timing as soon as *S* makes his choice.

Programming is accomplished by means of five rows of toggle switches and related

equipment, providing preprogramming for as many as 20 trials. Both absolute and dependent programming are possible. In absolute programming, the position of S_+ and the direction of any forced trial are programmed in advance. In dependent programming, on the other hand, either of these factors may depend on the outcome of any single previous trial. For example, one may program S_+ to be on the right on Trial 18 if *S* went left on Trial 2, or one may wish to force *S* on Trial 10 to the same (or opposite) side as his choice on Trial 7.

Spatial discrimination problems are programmed by having identical illumination occur in the two opposing arms. Finally, by always keeping one door closed, ALDA becomes an abbreviated runway.

An essential part of the programming panel is the visual display of *S*'s behavior, provided by a running-time counter and three pilot lights that simulate the maze. At the beginning of a trial the counter is reset to zero, and the light corresponding to the arm in which *S* is located is brightly lighted. As the doors open, the counter starts counting; and when *S* makes a choice, the counter stops and the light of the arm chosen goes full bright, while the light corresponding to the starting arm of that trial goes dim. Thus, *S*'s progress on every trial is easily followed from the visual display.

Multiple units of any automatic apparatus would not be feasible without some means of automatically recording the data. A Clary 11-column print-out machine is used for this purpose. The results of each trial are printed out as 11 digits, giving full information concerning the program on that trial and *S*'s choice and running time.

Each of the four Y mazes is separately housed in a sound-deadened cubicle, with white noise being fed down over each unit from a speaker in the ceiling. The laboratory, containing the vivarium and the experimental rooms, is maintained at 75° F. ± 5°, and the relative humidity is kept at 50% ± 15%. (In the present study, however, the humidifier was not yet operative and the relative humidity varied between 12% and 40%.)

Pretraining. Four days prior to discrimination learning, *S*s were placed on a 23-hr. water deprivation schedule, food being available ad libitum. The only pretraining given in the Y maze consisted in permitting *S* to drink eight dipperfuls of water in one arm of the maze with lights out and doors closed.

Discrimination training. For all *S*s, S_+ was the brightly lighted arm (10.0 ft.-c.), while S_- was the dimly lighted arm (0.15 ft.-c.). (These readings were taken under standard contitions, i.e., with the target placed on the grid floor, its leading edge 2.5 in. from the closed door.) The starting stimulus was a flashing light of moderate intensity. The intervals employed were: reward, 1.5 sec. ± .5 sec.; stimulus, 8 sec.; 57 sec. intertrial (lights out) ±3 sec.; and 4.5 sec. starting, ±.5 sec.

All *S*s were run 10 trials on Day 1 and 20 day thereafter until the criterion of 18 correct out of 20 successive trials was reached. The position of S_+ was determined by a standard program, 01001001110110001101, with "0" being interpreted as left and right on alternate days.

The average time of deprivation at the start of a day's trials was 22.2 hr. with a range of 18–25 hr.

Reversal learning. Upon reaching criterion each of the 30 *S*s was assigned to either Group R or Group ER, resulting in two groups of 15 *S*s each matched with respect to mean number of trials to criterion. The *S*s of Group R started on reversal learning the day after meeting criterion (i.e., the dimly lighted arm was now S_+), receiving 20 trials per day. The *S*s of Group ER, on the other hand, were first subjected to 60 extinction trials, 20 per day, and began reversal training on the fourth day. It should be noted that the extinction trials differed from the reversal trials only in that reward was absent in the former. For convenience, these extinction trials are referred to as unrewarded reversal trials and a choice of the dimly lighted arm is called a "reversal" response. Reversal training continued until the same criterion, 18 correct responses out of 20 trials, was reached. In addition, in order to obtain a reversal learning curve, all *S*s were run a minimum of 12 days (240 trials). The average deprivation was 22.2 hr.; range: 19–24.5 hr.

Results and Discussion

Discrimination. The matching procedure resulted in mean numbers of trials to criterion of 121.5 and 123.4 for Groups R and ER, respectively. Since discrimination and reversal learning scores were significantly correlated in Group R (rho = .604, $P = .05$), the importance of some matching procedure is indicated.

Reversal. Figure 2 summarizes the reversal data. Two things are clear from the figure. First, it shows that, as compared to Group R, Group ER made considerably more "reversal" responses on Days 1–3, during which time it will be recalled Group ER was on extinction. The differences are, by the t test, significant for Days 2 and 3 and for the totals of Days 1 through 3 (the respective Ps were .02, .01, and .001). The mean number of reversals for the total 60 trials was 8.8 for Group R and 14.3 for Group ER. Thus, removal of reward during the early trials of reversal learning serves to increase markedly responses to the former S–, fulfilling the premise of the hypothesis that initiated this study.

Second, however, it will be observed that as soon as reward is reinstituted, on Day 4, the performance of Group ER deteriorates, dropping below that of Group R and remaining there for the duration of reversal training. In fact, from Day 4 on the two graphs are roughly parallel with the curve of Group ER running at least two days behind that of Group R. The difference in the mean number of trials to reversal criterion reflects this

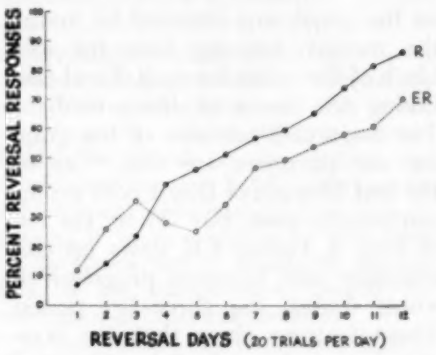


FIG. 2. Group learning curves during the first 12 days on reversal learning.

displacement, being 191.0 for Group R and 243.4 for Group ER (medians: 197.0 vs. 243.0). This difference is significant by both the t and the U tests ($P < .005$ in both cases). Thus, far from facilitating discrimination reversal learning, early extinction trials severely impede such learning.

Analysis of the role of reward during reversal learning. Analysis of the role of reward during reversal learning reveals that this factor serves at least three important functions.

First, consider the incentive function of reward. In Fig. 3 the running times are graphed for the first three days of reversal learning. Each point

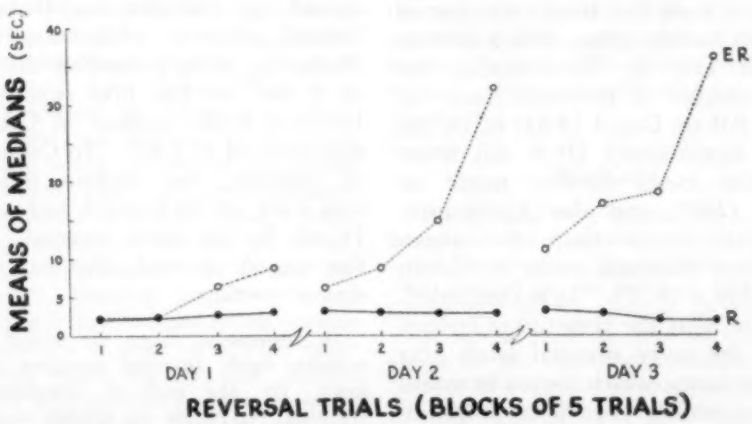


FIG. 3. Group running-time curves for the first 60 trials of reversal learning.

on the graph was obtained by finding the median running time for each block of five trials for each *S* and then taking the mean of these medians. The important features of the graph for our purposes are that: (*a*) for the first 10 trials of Day 1 both groups run equally fast; but (*b*) by the end of Day 1, Group ER slows up considerably and becomes progressively worse during the three-day period. These features show that the occasional reward received by Group R serves to keep *S*s of that group running rapidly albeit to the wrong side. The efficacy of the incentive function of reward is remarkable in view of the fact that during the first reversal day *S*s of Group R obtained only 1.33 reinforcements on the average.

Next, consider the cue function of reward. As noted earlier, on Days 2 and 3 (Fig. 2) Group R makes significantly fewer reversals than Group ER, a fact which is attributed to the presence of reward on the reversal responses of the former group.

A more striking demonstration of the cue effect of reward occurs in Group ER upon the reintroduction of reward on Day 4 of reversal learning (Fig. 2). As previously mentioned, introduction of reward on Day 4 results in a decline in the number of reversals in this group, which persists through Day 5. Statistically, the mean number of reversals made by Group ER on Day 4 (5.53) is, by the *t* test, significantly ($P = .05$) fewer than the mean number made on Day 3 (7.00), and also significantly ($P = .01$) fewer than the mean number of reversals made by Group R on Day 4 (8.27). It is concluded, therefore, that the presence of reward during the early reversal trials acts as a cue factor which serves to maintain responding that is appropriate to the initial discrimination.

The last function of reward to be considered is what may be called the *avoidance extinction function of reward.* As already seen, presence of reward during the first three reversal days leads to faster reversal learning than does absence of reward. On the other hand, the cue function of reward serves to keep *S* responding in accordance with the old discrimination. Nevertheless, something about the reversal problem is being learned, since these *S*s reverse faster than the *S*s of Group ER. The present hypothesis is that during these early reversals reward acts to extinguish *S*'s avoidance of the former S_-. The evidence for this makes use of the running-time scores of the reversal trials and of the immediately preceding trials of the first six reversals. (One *S* in Group ER made only five reversals during the first three reversal days, limiting to six the number of initial reversals available for comparative purposes.) For each of these reversals a running-time difference was obtained, i.e., the difference between the running time on that reversal and the running time of the immediately preceding trial. If reward acts to reduce the avoidance of the former S_-, then these differences should get smaller in Group R. Indeed, this is what was found. Beginning with a median difference of 4 sec. on the first reversal, by Reversal 6 the median of Group R was reduced to 1 sec. In Group ER, in contrast, the median difference was 3 sec. on Reversal 1 and reached 11 sec. by the sixth reversal. After the second reversal, the two curves do not overlap.

In summary, then, we would characterize early reversal learning as follows. By the end of discrimination training, S_+ builds up sizable approach tendencies and S_- develops rather strong

avoidance response tendencies. In order to reverse, *S* must extinguish in some measure the approach tendencies to the former S_+ but even more important, he must also extinguish the avoidance tendency toward the former S_-. Reward, aside from its incentive function, serves to keep *S* approaching the old S_+, and since these responses are unreinforced, they undergo extinction. Finally, the rewarded reversal trials act to extinguish avoidance of the former S_-. In our opinion this is by far the most important function. It will be recalled that the mean difference between the groups in arriving at the reversal criterion was of the order of 52 trials. Since in Group ER the 60 unrewarded reversal training trials could only serve to extinguish the approach tendencies to the former S_+, we attribute the eight trial saving to this factor. Conversely, the 52-trial advantage of Group R over Group ER may be attributed to the few (8.8 on the average) rewarded reversals that occurred during the first 60 reversal training trials. Later in reversal learning, other factors such as partially reinforced position responding, enter to complicate the picture and protract the course of reversal learning.

We wish, finally, to comment on our application of the notions of approach and avoidance to discrimination learning. Some years ago, Nissen (1950) proposed that discrimination learning could be handled in terms of approach-avoidance response categories, a proposal that promptly drew criticism from some quarters (Weise & Bitterman, 1951). In our view this idea has great merit, though we would not necessarily defend the position that *all* discrimination learning should be so interpreted. However, application of the notions of approach and avoidance to simple discrimination learning leads to many interesting deductions and offers promise of integration of this topic with other important areas, e.g., conflict, extinction, and perhaps others. For example, we have maintained that the development of avoidance tendencies toward S_- is an important factor in the establishment of a discrimination and, conversely, reversal learning critically requires the extinction of this avoidance. Any agent, therefore, that acts to reduce avoidance tendencies should impede original discrimination learning but facilitate reversal learning. One possible class of such agents, if we are willing to generalize on the properties of avoidance tendencies, is the tranquilizers, heavily investigated in other connections.

It now seems fairly well established (Pubols, 1956; Reid, 1953) that discrimination reversal learning is *facilitated* by overlearning on the original discrimination. Our explanation is that very few errors are made during the overlearning trials, and this virtual termination of experience with S_- leads to a reduction in the avoidance tendencies (via generalization of approach tendencies from S_+, if through no other mechanism) earlier developed toward this stimulus, therefore rendering reversal learning easier. Our clear prediction is that, if *S* were forced to have a reasonable number of experiences with S_- during the overlearning trials, reversal learning would be impeded rather than facilitated. Or, equivalently, overlearning should not facilitate reversal learning in a successive discrimination situation, since avoidance of S_- is maintained by the very nature of the situation. Even these few examples may serve to illustrate the potential fruitfulness of conceiving of at least some discrimination learning in terms of the development of approach and avoidance tendencies toward S_+ and S_-.

Summary

Thirsty rats were trained on a simple brightness discrimination problem in an automatic Y maze. The *S*s of Group R were then reversed on the problem, while *S*s of Group ER first underwent 60 extinction trials before going on reversal. Although Group ER was responding significantly more frequently to the former S_- by the end of the extinction trials (as compared to Group R at the end of its first 60 reversal trials), Group R learned the reversal considerably faster than Group ER.

Three functions of reward—incentive, cue, and avoidance extinction—were pointed out and discussed. Some implications of an approach-avoidance interpretation of discrimination learning were considered.

REFERENCES

DUFORT, R. H., GUTTMAN, N., & KIMBLE, G. A. One-trial discrimination reversal in the white rat. *J. comp. physiol. Psychol.*, 1954, **47**, 248–249.

KENDLER, H. H., & LACHMAN, R. Habit reversal as a function of schedule of reinforcement and drive strength. *J. exp. Psychol.*, 1958, **55**, 584–591.

NISSEN, H. W. Description of the learned response in discrimination behavior. *Psychol. Rev.*, 1950, **57**, 121–131.

PUBOLS, B. H., JR. The facilitation of visual and spatial discrimination reversal by overlearning. *J. comp. physiol. Psychol.*, 1956, **49**, 243–248.

REID, L. S. The development of noncontinuity behavior through continuity learning. *J. exp. Psychol.*, 1953, **46**, 107–112.

WEISE, P., & BITTERMAN, M. E. Response-selection in discrimination learning. *Psychol. Rev.*, 1951, **58**, 185–195.

(Received May 11, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

LEARNING AND PERFORMANCE AS FUNCTIONS OF RATION SIZE, HOURS OF PRIVATION, AND EFFORT REQUIREMENT [1]

GLEN D. JENSEN

Northwestern University

Presumably, the psychology of animal learning will eventually reach a point of development at which exact numerical values from empirical equations fitted to parametric studies of hours of food privation (and other variables) will enter into predictive equations. The ease of making accurate predictions will vary in direct proportion to the percentage of empirical relations that are invariant over a wide range of conditions. One purpose of the present paper is to determine whether or not the relation of hours of privation to maze performance is a relation which is invariant over a certain range of conditions; namely, feeding rats various sizes of daily ration during the week preceding testing.

Ramond, Carlton, and McAllister (1955) have shown that rats which were allowed to eat a relatively small amount before privation ran significantly faster than rats which ate a relatively large amount of food before privation even though privation time was constant for both groups. Their finding implies that detailed and accurate prediction will require a *set* of empirical curves relating running time to hours of privation (one curve for each size of ration). But the findings of Ramond et al. (1955) allow only of the conclusion that the asymptotes of such curves will vary. A more interesting question is whether or not the *shapes* (e.g., slopes) of such curves will vary with ration size. If the shape remains invariant, then the concept of "hours of privation" may still turn out to be a highly "significant" one, in Bergmann's (1957) terminology (i.e., one that enters into many empirical laws in a simple manner). However, the finding that the shape changes as a function of ration size would portend the need to construct an elaborate predictive apparatus to handle motivational variables. At a minimum, it would be necessary to table pairs of slope values and asymptote values for an equation relating performance to privation time, one pair for each size of ration.

To determine whether the shape of the empirical relation between hours of privation and instrumental learning is a function of ration size, a factorial design is mandatory. A significant interaction between hours of privation and ration size would affirm that such a relationship holds. However, lack of a significant interaction would not necessarily negate the relationship.

In the present study three sizes of daily ration (6, 8, and 10 gm.) were varied factorially with three levels of food privation (1, 5, and 22 hr.). It was hypothesized that 6-gm. *S*s would show faster running speeds than the 10-gm. *S*s, with the 8-gm. *S*s showing intermediate speeds. Such a finding would confirm Ramond

[1] This article is based on a PhD dissertation presented to the Graduate School, Northwestern University. The author is greatly indebted to W. F. Hill, under whose guidance this research was performed. This research was supported by Grant G-2672 from the National Science Foundation.

et al. (1955). Furthermore, it was hypothesized that ration size and privation time would interact, with the 6-gm. *S*s showing a more gradual slope than the 8-gm. and 10-gm. *S*s. That is, it was predicted that the (undernourished) 6-gm. *S*s would run almost as fast under 1 hr. of privation as under 22-hr. privation, whereas the 10-gm. *S*s would run much slower under 1-hr. than under 22-hr. privation.

An effort variable was included primarily to determine whether it interacts with hours of privation or ration size. If effort requirement interacts with hours of privation, then it should not be treated as a subtractive term from ${}_{S}E_{R}$. An attempt was also made to determine whether many closely spaced trials of a highly effortful nature would lead to depressed acquisition curves.

METHOD

Subjects and design. A total of 90 male albino rats of the Sprague-Dawley strain was assigned by table of random numbers to 45 experimental conditions. The design was a $5 \times 3 \times 3$ factorial with 2 *S*s per cell. Three sizes of daily ration were used (6, 8, and 10 gm.), so that there were 30 *S*s in each group defined by daily ration size. There were three conditions of privation (1, 5, and 22 hr.), so that 30 *S*s underwent each condition. Finally, there were five conditions of effort requirement so that 18 *S*s were run under each of the five effort requirements. The definition of the various levels of effort requirement is treated below.

Because the *S*s came in several shipments from a stock farm, and because an analysis of variance on a previous study yielded a significant *F* between shipments, the random assignment was restricted in the following manner. Of the first 9 *S*s run, 3 were assigned to each of the three deprivation conditions and 3 were assigned to each of the three ration-size conditions. Of the first 10 *S*s run, 2 were assigned to each of the five effort-requirement conditions.

The *S*s ranged in age from 61 to 90 days with a mean of 77.9 days. In addition to the 90 *S*s which completed the experiment, three *S*s were lost for the following reasons: two *S*s failed to learn and one *S* was discarded because of a procedural error.

Apparatus. A T maze described by Cotton, Lewis, and Metzger (1958) was used with one modification: It was attached to a stand which allowed rapid changes of the angle of incline of the runway section of the maze. Two electric timers were automatically started when *S* put weight on a treadle located 2 in. from the startbox in the stem (runway) of the maze. One timer was stopped when *S* put weight on a treadle located 4 in. anterior to the choice point. The second timer was stopped when *S* touched a treadle located 3 in. from either goalbox. Arm doors prevented retracing. Time measurements were taken in .01 sec. A stopwatch was used to record the time lapse between the touch of *S*'s feet on the startbox floor and activation of the timers (the timers were activated when *S* touched the first treadle just inside the stem). This stopwatch measure will henceforth be called the "latency" measure. The values read from the two electric timers will be called "stem time" and "arm time," respectively.

Procedure. Prior to the first of seven days of "formal" pre-experimental food scheduling, *S*s were maintained in individual home cages for at least a week, during which time they were weighed at least once every two days and fed once a day. The size of the daily food ration varied with each *S* and was calculated to get each *S* as near as possible to 210 gm. of body weight by the first day of formal food scheduling. In almost all cases the *S* weighed less than 210 gm. before this initial treatment. On the first day of formal pre-experimental food scheduling, each *S* which was to be given acquisition training under the 22-hr. drive, 10-gm. maintenance condition was presented 10 gm. (±.2 gm.) of finely ground Purina laboratory chow in a plastic cup inside its home cage. Each *S* in this condition received its food 1 hr. after the time of day that it was scheduled to receive experimental treatment. From the treatment of this group, the reader can infer the treatment of the other two 22-hr. drive groups (those maintained on 8 and 6 gm. of daily ration). Regardless of ration size, *S* usually finished the ration within 20 min. Spillage was never enough to register on a scale calibrated to 0.1 gm. (This feeding procedure was chosen after extensive pilot work.)

The only difference in the treatment of the three 5-hr. drive groups from that of the corresponding three 22-hr. drive groups concerns the time of day at which *S*s were

fed, and similarly for the three 1-hr. drive groups.

Each of the nine groups (10 *S*s per group) representing combinations of the three levels of daily ration and the three levels of privation time was subdivided into five groups representing five levels of effort requirement. Effort requirement was varied by changing the incline of the stem. One group was run with the maze in a horizontal position. The other groups were run with the maze tilted at angles of 2.93°, 5.80°, 8.67°, and 11.50° upwards from the startbox.

During each of the seven days of formal food scheduling prior to the first experimental day (Day 8), each *S* was handled for 5 min. a day. During the 5 min., the *S* was allowed to eat six 45-mg. food pellets from a small cup which was later to be a goalbox feeder. If *S* did not eat all of the 6 pellets during the 5 min. of taming, the pellets not eaten were placed in his home cage.

On Day 8 each of the 90 *S*s received 30 acquisition trials, each "correct" run being rewarded with a 45-mg. food pellet. The rewarded, or correct, side was assigned by table of random numbers with 45 *S*s being rewarded on each side of the maze. On each trial *S* remained approximately 10 sec. in the goalbox, after which it was placed for 15 sec. in its living cage to await the next trial. Water was available in the living cage at all times. During the first five trials, all *S*s were run with the maze in the horizontal position. The mean speed for each *S* on these first trials was later used as a predictor in obtaining adjusted means for the "effort" groups at the end of acquisition. During the remaining 25 trials on Day 8 the maze was tilted or not tilted depending upon which condition *S* was in. Usually about 10 extra sec. elapsed between Trials 5 and 6 relative to any other trials.

On Day 9 each *S* received 15 acquisition trials under the same effort requirement that obtained during Day 8. Then, after a 10-sec. delay while the maze angle was reset, each *S* received an additional 15 acquisition trials with the maze in a horizontal position. (If the effort variable had caused depressed acquisition curves for the high-effort groups, then the acquisition curves of these groups should have approached the curve of the low-effort (horizontal) group during the last 15 trials of Day 9.)

During any acquisition trial, if *S* stayed in the startbox for 2 min. without activating the timer, or if *S* did not reach a goalbox within 2 min. after activation of the timers, it was manually lifted and placed in the "correct" goalbox and allowed to eat.

Each running-time value was converted to its reciprocal to meet the assumptions of analysis of variance. All statistical testing was done on these reciprocal scores.

Results

Ration size. With respect to the latency, stem time, and arm time measures, the ration-size groups appeared to approach different asymptotes. In each case the 6-gm. group approached the highest asymptotic response strength, while the 10-gm. group was always lowest, with the 8-gm. group intermediate. In the case of the stem time data and the latency data, the 10-gm. group showed superior response strength to the 8-gm. group during Trials 1–30. It is possible that the 10-gm. group was of superior initial ability on these response measures and that the ration-size variable gradually offset the difference. It was decided to test the effect of ration size by using the mean score of each *S* for Trials 31–60. If it is true that the 10-gm. group was of superior initial ability, the tests performed on Trials 31–60 should give conservative estimates of the effect of ration size. Analyses of variance were performed on the reciprocal stem time and reciprocal arm time data because the assumptions of normality of distribution and homogeneity of variance were not violated. The F between groups for the stem time data was 3.32 ($P < .05$, 2/87 *df*), and the F for the arm time data was 7.91 ($P < .01$, 2/87 *df*). A Kruskal-Wallis analysis of variance by ranks (1952) was performed on the latency data because these data were positively skewed. The H obtained was 12.07 ($P < .005$).

Since all three of the above-mentioned dependent variables show the

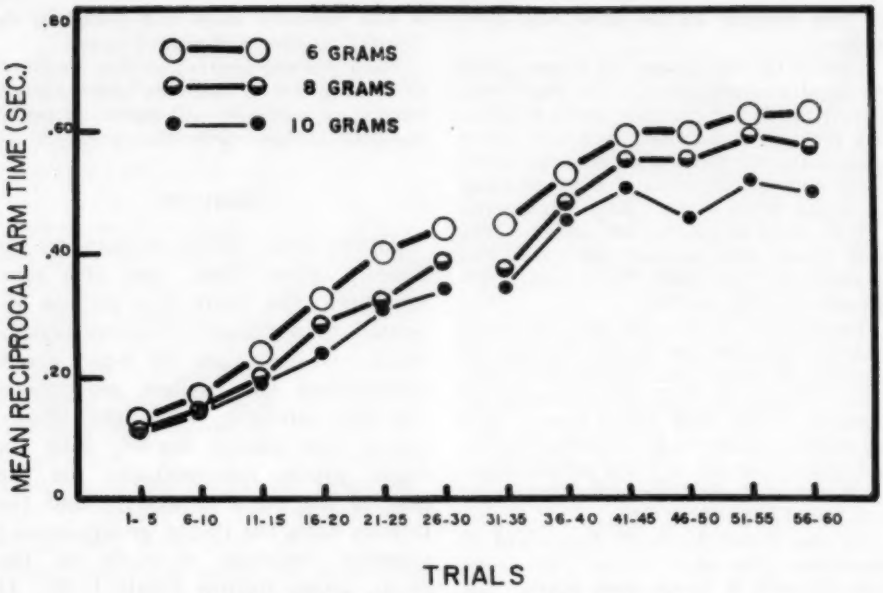


FIG. 1. Reciprocal arm time for ration-size groups.

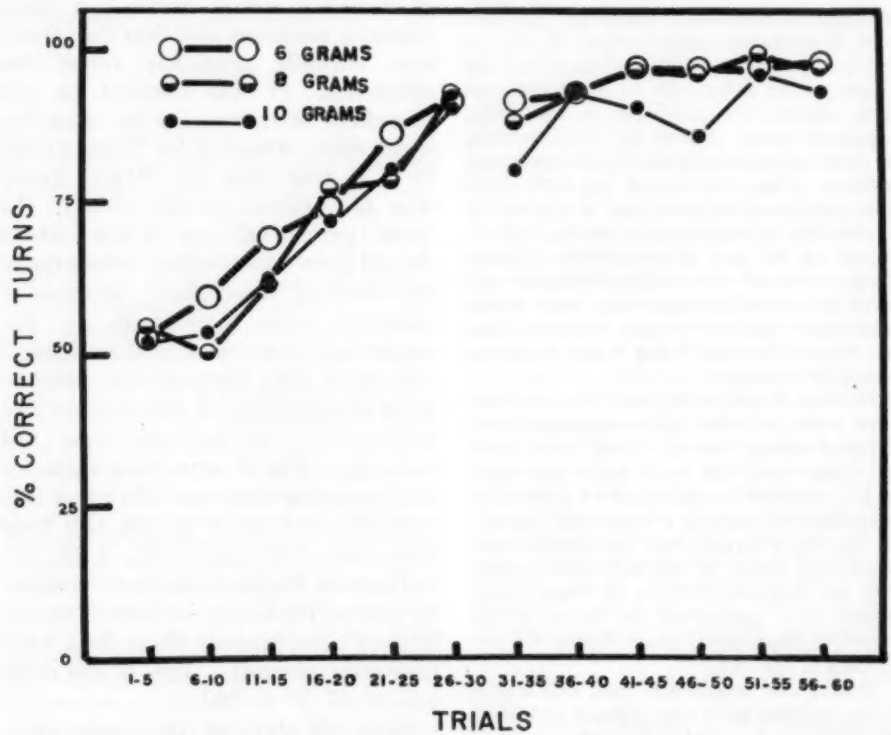


FIG. 2. Percentage of correct turns for ration-size groups.

TABLE 1

WEIGHT AND AGE DATA ACCORDING TO RATION-SIZE GROUPS

	Daily Ration Size		
	6 Gm.	8 Gm.	10 Gm.
Weight range, Day 1	207–218 gm.	208–217 gm.	205–219 gm.
Mean weight, Day 1	210.9 gm.	212.1 gm.	210.9 gm.
Weight range, Day 8	176–195 gm.	189–209 gm.	203–218 gm.
Mean weight, Day 8	183.7 gm.	197.4 gm.	211.9 gm.
Mean weight loss, Day 1–Day 8	27.2 gm.	14.7 gm.	−1.0 gm.
Age range, Day 8	61–90 days	61–90 days	61–90 days
Mean age, Day 8	77.7 days	77.3 days	78.7 days

Note.—Day 1 is the first of seven days of "formal" food scheduling. Day 8 is the first experimental day.

same general trend, only one (arm time) will be presented graphically (Fig. 1). Mean reciprocal arm time during Trials 31–60 correlated .92 with mean reciprocal stem time and −.94 with mean latency during the same trials.

The turn data (Fig. 2) do not fit the pattern of the other response measures. Note that on Trials 31–60 there is little difference between the 6-gm. and 8-gm. groups. The 10-gm. group appears, however, to be inferior to the other groups. Note that the largest disparity between the 10-gm. group and the other groups is at the data point representing Trials 46–50. Inspection of individual data revealed that most of this disparity was contributed by two *S*s in the 10-gm. group who made markedly more errors (relative to their own previous data) following the change in maze angle which occurred just before Trial 45. Therefore, in testing the effect of ration size, it appeared most conservative to exclude Trials 46–60. An analysis of variance was performed on the percentage-correct scores of each *S* for Trials 31–45. The *F* value obtained was 3.67 ($P < .05$, 2/87 *df*). Because the distribution of scores was not normal, a Kruskal-Wallis test was used to supplement the results of the *F* test. An *H* of 8.07 was obtained ($P < .02$).

The weight and age data of the *S*s broken down according to ration-size groups is shown in Table 1. Of special interest (for reasons to be presented below) is the fact that the 6-gm. *S*s lost an average of only 27.2 gm.

Hours of privation. Table 2 shows mean performance during Trials 31–60 as a function of privation time (except for the turn measure). Note that for the first three dependent variables (which are highly correlated) response strength is a direct function of privation time, whereas the percentage-correct shows no such relation. Tests exactly analogous to those used on the ration-size variable were performed on these data. Privation time had a

TABLE 2

MEAN PERFORMANCE DURING TRIALS 31–60 AS A FUNCTION OF PRIVATION TIME

Response Measure	Privation Time		
	1 Hr.	5 Hr.	22 Hr.
Reciprocal arm time	.485	.505	.567
Reciprocal stem time	.869	.916	.994
Latency	3.362	2.650	2.374
Percentage of correct turns (Trials 31–45)	91.4	90.6	91.1

TABLE 3

MEAN PERFORMANCE AS A FUNCTION OF MAZE INCLINE

Response Measure	Angle of Incline				
	0°	2.93°	5.80°	8.67°	11.50°
Reciprocal arm time (Trials 1–45)	.322	.325	.330	.347	.300
Reciprocal stem time (Trials 1–45)	.725	.712	.705	.713	.641
Latency (Trials 31–45)	2.15	3.22	2.29	2.80	3.27
Percentage of correct turns (Trials 31–45)	91.8	88.5	90.7	91.8	92.6

significant effect ($P < .06$ or better) on each of the first three response measures, but not on the turn measure.

Effort. Table 3 shows mean performance as a function of maze incline. Note that none of the response measures showed an orderly ranking of means. Significance tests for the first two response measures gave no suggestion of a main effect of effort requirement (maze incline). The other data, being (graphically) even less suggestive of a main effect, were not analysed.

TABLE 4

SUMMARY OF ANALYSIS OF VARIANCE OF RECIPROCAL ARM TIME SCORES

Source	*df*	*MS*	*F*
Ration size (R)	2	9.528	3.24*
Hours of privation (H)	2	11.374	3.87*
Effort (E)	4	1.133	0.39
R × H	4	1.027	0.35
R × E	8	3.039	1.04
H × E	8	1.312	0.45
H × R × E	16	1.183	0.40
Between *Ss* in same group	45	2.935	
Trials (T)	8	49.868	267.53**
T × R	16	0.235	1.26
T × H	16	0.247	1.33
T × E	32	0.191	1.02
T × R × H	32	0.229	1.23
T × R × E	64	0.085	0.45
T × H × E	64	0.116	0.62
T × R × H × E	128	0.157	0.84
Pooled *Ss* times trials	360	0.186	
Total	809		

* $P = .05$.
** $P = .01$.

Interactions. The Kogan (1948) analysis of variance for trend, extended to handle three independent variables of classification, appeared to be the most efficient way to test for all possible interactions. Two identical tests were employed, one on the reciprocal stem time data and one on the reciprocal arm time data. In order to avoid a bias in the selection of trend points, it was decided to use all of the data obtained while all three variables were operative. This necessitated the exclusion of all trials after Trial 45. After Trial 45 the effort variable was no longer operative: so inclusion of such trials would create a bias in favor of a significant interaction of effort and trials. In each analysis nine trend points were used. Each trend point was a mean of a block of five trials. Not one of the 22 *F* ratios for interaction was significant or even suggestive. Sixteen of the *F* ratios were below 1.0 and the highest was 1.3.

Table 4 shows the results of the analysis of the arm time data. The table is presented to illustrate the technique used as well as the results obtained.

The turn data were used to test the interaction of the independent variables. The percentage-correct scores for Trials 31–45 entered into a 3×3×5 analysis of variance. Again, none

of the interaction terms approached significance.

Discussion

Ration size: Main effect. The evidence is clear-cut that ration size is inversely related to arm time, stem time, and latency. Thus the findings of Ramond et al. (1955) are confirmed and extended. One implication is that the asymptote of an empirical equation fitted to a parametric study of hours of privation will be a function of ration size, if any of these response measures are used.

The fact that the ration-size variable produced rather large differences in performance even though the range of ration sizes was not great (nor the resulting weight changes) indicates that it is advantageous to control this variable in animal studies in general. Control of this variable should reduce within-groups variance. Control can be greatly increased by using the feeding procedure herein employed (i.e., using a constant amount of ground food rather than a procedure which requires skill on the part of the rat to ingest a large percentage of the ration or to ingest a large quantity in a unit of time).

The present study offers some evidence that choice behavior is also a function of ration size, but the relationship is far from clear.

Interaction of ration size and hours of privation. The lack of significant interaction between ration size and hours of privation is an important and fortunate finding. The relation of hours of privation to performance has shown an invariance (in general shape of the function) over a range of conditions within which it might reasonably be expected to vary. Such invariances are highly desirable in any science because their presence makes the task of prediction much easier.

If the shape of the relation of hours of privation to performance had been shown to be a function of ration size (by a significant interaction of ration size and privation time), then the existing parametric studies of privation time would be of much less value to "future" psychologists (those attempting precise and detailed predictions of performance using exact numerical values for all relevant variables). This is true because in those studies the feeding procedure is such that it is not possible to state the amount that *S* actually ate per day.

It should be cautioned that failure to reject the null hypothesis does not confirm it. There may indeed be an interaction of ration size and hours of privation. However, if such an interaction exists, it is probably of small consequence or the interaction *F* ratio would have been above 1.0.

It should also be noted that the present study shows only that no significant interaction occurred when ration size was varied over the range of 6–10 gm. for 7 days. Presumably there is a value of ration size so low that all *S*s (regardless of privation time) would be sick or dead, so that a spurious interaction would be incurred by its use. The 10-gm. *S*s in the present study appeared to be quite healthy and gained slightly in weight, while the 6-gm. *S*s were (subjectively) quite thin by the time of experimentation. Hence, it is possible to conclude that within a reasonable range of ration size there is no interaction with privation time.

Effort: Main effect. The fact that the effort variable produced no acquisition main effect in this study places it at odds with three recent studies (Aiken, 1957; Appelzweig, 1951; Montgomery, 1951). The reason for this disparity may lie in the difference in response measures employed. Montgomery and Appelzweig may have obtained a relation between bar pressing and effort because a rat must get into a better position to make a highly effortful press than to make a light press, and time elapses while position is improved. Aiken's group differences could represent pain tolerance effects rather than fatigue or I_R effects because (presumably, at least) pushing a panel with the nose is painful.

The present study demonstrates that within a fairly wide range, effort is not

an important variable in maze behavior. This is a particularly fortunate finding because there is a sense in which the effort variable is confounded in every maze experiment that employs a vigor measure. Any variable that increases the vigor of a response also increases the effortfulness of the response (the greater the vigor, the greater the kinetic energy gained, hence the greater the effort expended). The present study indicates that such confounding has probably not affected the validity of experimental claims to any serious degree.

Interaction of effort and trials. Solomon (1948) has pointed out that Hullian work-inhibition theory leads to a clear-cut prediction that if effortfulness of a response is increased while intertrial interval remains constant (and short), then depressed acquisition curves will result. This hypothesis was tested but not confirmed by the F ratio for interaction of trials and effort in two analyses. It would be interesting to see whether depressed acquisition curves can be produced by using a longer maze (constant for all groups) so that more potential energy would have to be gained on any trial for a specific angle of maze incline. It is not recommended that maze angles greater than 12° be employed because of slippage which spuriously increases running time. However, the present study does not permit of an optimistic forecast for such a research effort. Rather, the present study indicates that work-inhibition theory needs revision.

Summary

Ninety rats were given 60 rewarded trials in a T maze. The design was a $5 \times 3 \times 3$ factorial with 2 *Ss* per cell. Three sizes of daily food ration (6, 8, and 10 gm.) were varied over a 1-week period. There were three conditions of food privation (1, 5, and 22 hr.), and five conditions of effort requirement defined in terms of angle of maze incline (0°, 2.93°, 5.80°, 8.67°, and 11.5°).

The ration-size and hours-of-privation variables both produced significant effects upon acquisition running speed and latency of running. The ration-size variable produced a significant effect upon choice behavior, but the hours-of-privation variable did not. The effort-requirement variable gave no suggestion of a main effect. No significant interaction effects were found.

References

Aiken, E. G. The effort variable in the acquisition, extinction, and spontaneous recovery of an instrumental response. *J. exp. Psychol.*, 1957, **53**, 47–51.

Appelzweig, M. H. Response potential as a function of effort. *J. comp. physiol. Psychol.*, 1951, **44**, 225–235.

Bergmann, G. *Philosophy of science.* Madison: Univer. Wisconsin Press, 1957,

Cotton, J. W., Lewis, D. J., & Metzger, R. Running behavior as a function of apparatus and of restriction of goal box activity. *J. comp. physiol. Psychol.*, 1958, **51**, 336–341.

Kogan, L. S. Analysis of variance-repeated measurement. *Psychol. Bull.*, 1948, **45**, 131–143.

Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Ass.*, 1952, **47**, 583–621.

Montgomery, K. C. An experimental investigation of reactive inhibition and conditioned inhibition. *J. exp. Psychol.*, 1951, **41**, 39–51.

Ramond, C. K., Carlton, P. L., & McAllister, W. R. Feeding method, body weight, and performance in instrumental learning. *J. comp. physiol. Psychol.*, 1955, **48**, 294–298.

Solomon, R. L. The influence of work on behavior. *Psychol. Bull.*, 1948, **45**, 1–40.

(Received May 13, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

STIMULUS MEANING IN STIMULUS PREDIFFERENTIATION [1]

SHEILA M. PFAFFLIN [2]

Johns Hopkins University

It has been found that pretraining with stimuli will facilitate subsequent learning of responses to those stimuli, even though the pretraining was not of such a nature that transfer would have been predicted from response generalization (Baldwin, 1954; Cantor, 1955; Foster, 1953; Gagné & Baker, 1950; Goss, 1953). Such facilitation has come to be called stimulus predifferentiation, and no explanation for it has yet been established (Arnoult, 1957; Vanderplas, 1958). The hypothesis by Miller and Dollard (1950) that labels learned as responses to stimuli subsequently act as cues which make the stimuli more distinctive, thus decreasing stimulus generalization and facilitating learning, has been advocated most often for this purpose. Unfortunately, this distinctiveness-of-cues hypothesis has difficulty in explaining the results of several experiments on stimulus predifferentiation. Highly distinctive labels did not produce more facilitation than similar labels when used as responses during pretraining (Gerjuoy, 1953). It is not even necessary to have a different label for every stimulus to obtain facilitation (Hake & Eriksen, 1955; Robinson, 1955). Yet the hypothesis would actually predict a decrement in subsequent performance due to acquired similarity when many stimuli had the same label.

Effort has therefore been directed to the study of additional variables. The one most frequently studied has been meaning, usually measured by number of verbal associations in the manner suggested by Noble (1952). The results of these studies have generally been discouraging (Dysinger, 1953; Goss & Greenfeld, 1958). Arnoult, however, found that if the label used during pretraining seems to *S*s to be appropriate for the stimulus, to "belong" to it, it may especially facilitate subsequent ability to recognize the stimulus. It would appear that perhaps it is not the meaningfulness of the labels which is crucial for stimulus predifferentiation, but rather the meaning of the label in relation to the meaning of the stimulus.

In the present experiment the stimuli (visual forms) were selected by a new criterion of meaningfulness; the consistency with which *S*s supplied the same label for the stimulus when asked to name it. The reasoning was that stimuli which had very well-defined meanings would receive the same label from most *S*s, while stimuli without well-defined meanings would not be given any single label with high frequency.

The modal label assigned by *S*s to a stimulus, regardless of its actual

[1] This paper is based on a dissertation submitted in partial fulfillment of the requirements of the PhD degree at Johns Hopkins University. The author is greatly indebted to Clinton B. DeSoto, under whose supervision this study was conducted, for his guidance and helpful criticism, and to W. R. Garner for his advice in the planning of this experiment and in the preparation of this paper. The author is also indebted to James Deese for his encouragement and advice during the early phases of this experiment.

[2] Now at Bell Telephone Laboratories, Murray Hill, New Jersey.

frequency, was considered to be the most appropriate or relevant label for that stimulus.

Once the meaningfulness of the stimuli and their relevant labels had been determined in this way, a stimulus predifferentiation experiment was conducted to investigate the following three questions. Is *S*'s ability to learn responses to stimuli without pretraining related to the meaningfulness of the stimuli? How is learning influenced by the following types of pretraining: learning of relevant labels for stimuli; learning of irrelevant labels for stimuli; observation of the stimuli? Do the effects of different kinds of pretraining vary for stimuli which differ in meaningfulness?

Method

Selection of stimuli. Visual forms were used as stimuli since *Ss* are accustomed to labeling stimuli of this sort. From an initial pool of 85 silhouette forms, 30 forms were selected for the stimulus predifferentiation experiment. The initial pool was drawn from various sources so as to provide a set that would be expected to vary widely in meaningfulness. Many of them were adapted from the nonsense forms used by Arnoult (1956b).

The forms were cut out of black construction paper and mounted on heavy white paper. They had a maximum diameter of 6 in. In the procedure used to determine consistency of labeling, each form was exposed to *Ss* for 20 sec. The *Ss* were instructed to write the name of each form as it appeared. The *Ss*, 50 male students randomly drawn from the introductory psychology course at Johns Hopkins University, were run in groups not exceeding 10. The forms were presented in a different random order to each group of *Ss*

The modal name for each form was obtained by determining the noun most frequently given in response to it. The consistency with which each form was labeled was defined as the percentage of the responses to that form which used the modal label. Three sets of forms were then selected for use in the main experiment. They consisted of those 10 forms which were labeled most consistently (average consistency = 94%); those 10 forms which were labeled with medium consistency (average consistency = 48%); and those 10 forms which were labeled with lowest consistency (average consistency = 12%). These forms are shown in Fig. 1 with their modal labels.

Pretraining with the stimuli. The kinds of pretraining used were: observation; the learning of irrelevant labels; the learning of relevant labels; and no pretraining. These four kinds of pretraining were made orthogonal to the three form classes so that 12 conditions resulted. The modal label given each form was used as its relevant label. The following 10 adjectives were used as the irrelevant labels for all three classes of forms: fresh, true, slow, better, complete, proper, weak, loud, alone, late.

The forms were photographed and made into 35-mm. slides. Labels appeared under the forms on those slides used in learning of the labels. The apparatus consisted of an Argus slide projector, a beaded screen, and a Timolite timer.

The *Ss* were 120 male students randomly drawn from the introductory psychology course at Johns Hopkins University. None of them had served as *Ss* in selecting the stimuli. Ten *Ss* were randomly assigned to each experimental condition. The *Ss* in the three control conditions received no pretraining. The remaining 90 *Ss* received pretraining with the stimuli under the following procedure. They were run in groups of not more than four. Those *Ss* who learned labels as responses to the forms saw the forms 12 times: 6 times with labels and 6 without, alternately. On the trials when the labels were not present, *Ss* were required to write down the label for each form. Those *Ss* who were to observe only were shown the forms 12 times without labels. The forms were presented for 2 sec. each under all conditions. Different random orders of presentation of the stimuli were used throughout.

Discrimination learning. All 120 *Ss* then learned the discrimination task under the same conditions. The *Ss* were given response boards which had two buttons on them. They were told that they were to learn which button to press in response to each of the forms which they would be shown. The forms were then presented to them without labels by means of an Argus projector. They were given 12 learning trials, and the order in which the forms were presented was randomized within each trial. The forms appeared on the screen for 4 sec. Following

FIG. 1. The forms used as stimuli in this experiment, grouped according to whether they were labeled with high, medium, or low consistency.

each presentation of the form, *E* told *Ss* which button should have been pressed. One button was correct for five of the forms, and the other button for the other five forms. The *Ss*' responses were recorded by an Esterline-Angus 20-pen recorder.

RESULTS

The number of errors made by each *S* during the last 11 learning trials was determined. The first trial was excluded from this calculation since *Ss* were necessarily guessing on it. Mean errors for the 12 experimental groups are given in Table 1. In general, the greatest number of errors occurred in responses to forms labeled with low consistency; the fewest in responses to forms labeled with high consistency. This relationship is reversed following pretraining by learn-

TABLE 1

MEAN NUMBER OF ERRORS DURING DISCRIMINATION LEARNING FOR ALL CONDITIONS

Pretraining	Form Classes			Means
	High	Medium	Low	
None	11.1	22.1	30.1	21.1
Observation	6.6	13.1	19.0	12.9
Irrelevant Labels	13.9	17.1	19.1	16.7
Relevant Labels	15.8	15.1	12.8	14.6
Means	11.8	16.8	20.2	16.3

ing of relevant labels. All kinds of pretraining result in a decrease in errors in comparison with the number made when no pretraining is given.

Analysis of variance of the error scores shows that both form classes and kinds of pretraining were significant sources of variance (Table 2). When the decreases in errors following pretraining were tested for significance by t test, they were found to be significant only if pretraining had been by observation or by learning of relevant labels. In view of the apparently linear relationship between consistency of labeling and number of errors, the differences between form classes were not tested for significance by t test. The following method was used instead.

The three form classes are approximately equally spaced on a metric, the consistency measure. The usual analysis of variance does not take this fact into account in its partition of the variance. It is possible to do so, however, by determining how much of the variance is accounted for by the linear correlation between the error scores and the form classes. The Pearson r was calculated by taking the intervals between form classes to be equal, and assigning the form classes the arbitrary values 1, 2, and 3. This procedure was considered to be justified since the average consistency values for the form classes depart only slightly from equal spacing, and these values were themselves subject to some chance factors in their determination. The linear component of the variance due to form classes was then calculated and found to be significant at the .01 level. There is an inverse relationship between the amount of consistency shown in labeling forms and the number of errors made in learning to respond to them. The nonlinear component is not significant, and is in fact less than would be expected from chance variation.

It is also possible to separate the interaction into linear and nonlinear components by applying the same procedure to the scores for the different kinds of pretraining separately. The linear component is significant at the .05 level, while the nonlinear component is again unexpectedly small. The entire interaction is not significant at the .05 level, though it approaches this level. Examination of the linear trend for each kind of pretraining separately shows that with

TABLE 2

ANALYSIS OF VARIANCE OF THE ERRORS MADE DURING DISCRIMINATION LEARNING

Source	*df*	*MS*	*F*
Form Classes (F)	2	738.03	6.46**
Linear Component	1	1462.03	12.88**
Nonlinear Component	1	14.04	
Predifferentiation Type (P)	3	379.46	3.32*
F × P Interaction	6	221.90	1.94
Linear Interaction	3	441.65	3.86*
Nonlinear Interaction	3	2.15	
Within Cells	108	114.31	

* $P = .05$.
** $P = .01$.

no pretraining, or with observation only, there is a significant linear component. The linear component is not significant for either of the labeling conditions.

Discussion

The results of this experiment support the view that both stimulus meaning and response meaning are important factors in stimulus predifferentiation. Of particular interest is the finding that stimuli which differ in meaningfulness are affected differently by various kinds of stimulus predifferentiation. Learning responses to forms low in meaning seems to be especially facilitated by pretraining with relevant labels. Learning responses to forms high in meaning is apparently facilitated by observation as much as by any other kind of pretraining, and may actually be hindered by pretraining with relevant labels. The distinctiveness-of-cues hypothesis as now interpreted seems inadequate to explain this finding, since it provides no reason why relevant labels should be especially distinctive when learned as responses to forms low in meaning. A supplementary explanation is suggested by the "attention-to-cues" hypothesis favored by Hake and Eriksen (1955). They believe that stimulus predifferentiation results from *S*s' learning to notice the distinctive features of the stimuli. The fact that responses are learned during paired-associate pretraining is ordinarily incidental, the important aspect of such pretraining being that it forces *S*s to observe the stimuli carefully. It would follow from this view that relevant labels, if they help *S*s to select distinctive features of the stimuli, would aid subsequent learning.

The problems remain of how relevant labels help *S*s select relevant features of the stimuli, and why they do so only for stimuli low in meaning. To answer these questions it is necessary to make some assumptions about what constitutes distinctiveness. Let us suppose that people build up concepts, categories, or central processes from their past experience, and that these concepts include one or more labels. Further, let us suppose that *S*s are aided in discriminating among future stimuli by relating them to these categories. Stimuli are distinctive which are easily related to different categories. Such stimuli are also meaningful. Stimuli low in meaning are hard to relate to any category. In the absence of appropriate categories, *S*s, by observing the stimuli repeatedly, may develop new categories which will aid them in discriminating among the stimuli, but this would necessarily be slower than using already established categories. In the case of stimuli low in meaning, the relevant labels suggest appropriate categories, thus making it easier for *S*s to discriminate whatever distinctive features the stimuli have. Irrelevant labels would not have this additional selective effect, and contribute to stimulus predifferentiation only by forcing *S*s to observe the stimuli carefully. Stimuli high in meaning can be easily related to a concept without the help of relevant labels, and subsequent learning might actually be hindered by relevant labels if their repeated use induced fatigue and made the concept harder to activate. This approach seems consistent with views of the role of central processes in learning and perception that have been advocated by others (Bartlett, 1932; Hebb, 1949; Humphrey, 1951). It is also consistent with the original interpretation of the distinctiveness-of-cues hypothesis in which the term responses could include central processes associated with the stimuli. At the same time it explains why labels arbitrarily chosen as responses for stimuli without regard for the meaning of the stimuli have not produced the induced similarity or distinctiveness expected from the properties of the labels alone.

Summary

Three sets of 10 visual forms each were selected so that the sets differed in meaningfulness. Meaningfulness was measured by the consistency with which 50 *S*s labeled each form. Another 120 *S*s were then

randomly assigned to one of the sets of forms. These *Ss* were required either to observe the forms, to learn irrelevant labels for the forms, to learn appropriate labels for the forms, or were given no pretraining with the forms. Next these *Ss* learned to press one of two buttons in response to each form in their set.

The results show an inverse linear relationship between number of errors made in learning the button pressing response and meaningfulness of the stimuli. Both observation and learning of labels appropriate to the stimuli produced significant decreases in errors below the number found without pretraining, but the linear relationship did not hold after learning of appropriate labels. The results are interpreted to support the view that the meaning of both stimuli and responses used during predifferentiation training must be considered in relation to each other before the process can be adequately understood.

REFERENCES

ARNOULT, M. D. Recognition of shapes following paired-associates pretraining. In G. Finch & F. Cameron (Eds.) *Symposium on Air Force human engineering, personnel, and training research.* Washington, D. C.: Nat. Acad. Sci.—Nat. Res. Coun., Publ. 455, 1956, 1–9. (a)

ARNOULT, M. D. Familiarity and recognition of nonsense shapes. *J. exp. Psychol.*, 1956, **51**, 269–276. (b)

ARNOULT, M. D. Stimulus predifferentiation: Some generalizations and hypotheses. *Psychol. Bull.*, 1957, **54**, 339–350.

BALDWIN, R. D. Discrimination learning as a function of stimulus predifferentiation and mediated association training. Unpublished doctoral dissertation, State Univer. of Iowa, 1954.

BARTLETT, F. C. *Remembering: A study in experimental and social psychology.* Cambridge, Eng.: Cambridge Univer. Press, 1932.

CANTOR, J. H. Amount of pretraining as a factor in stimulus predifferentiation and performance set. *J. exp. Psychol.*, 1955, **50**, 180–184.

DYSINGER, D. W. An investigation of stimulus predifferentiation in a choice discrimination problem. Unpublished doctoral dissertation, State Univer. of Iowa, 1951.

FOSTER, H. W. Stimulus predifferentiation in transfer of training. Unpublished doctoral dissertation, Univer. of Michigan, 1953.

GAGNÉ, R. M., & BAKER, K. Stimulus predifferentiation as a factor in transfer of training. *J. exp. Psychol.*, 1950, **40**, 439–451.

GERJUOY, I. R. Discrimination learning as a function of the similarity of the stimulus names. Unpublished doctoral dissertation, State Univer. of Iowa, 1953.

GOSS, A. E. Transfer as a function of type and amount of preliminary experience with task stimuli. *J. exp. Psychol.*, 1953, **46**, 419–427.

GOSS, A. E., & GREENFELD, N. Transfer to a motor task as influenced by conditions and degree of prior discrimination training. *J. exp. Psychol.*, 1958, **55**, 258–269.

HAKE, H. W., & ERIKSEN, C. W. Effect of number of permissible response categories on learning of a constant number of visual stimuli. *J. exp. Psychol.*, 1955, **50**, 161–167.

HEBB, D. O. *Organization of behavior.* New York: Wiley, 1949.

HUMPHREY, G. *Thinking.* New York: Wiley, 1951.

MILLER, N. E., & DOLLARD, J. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.

NOBLE, C. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.

ROBINSON, J. S. Effect of learning verbal labels for stimuli on their later discrimination. *J. exp. Psychol.*, 1955, **49**, 112–114.

VANDERPLAS, J. M. Transfer of training and its relation to perceptual learning and recognition. *Psychol. Rev.*, 1958, **65**, 375–385.

(Received May 21, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

SUPPLEMENTARY REPORT: PROMPTING VERSUS CONFIRMATION IN PAIRED-ASSOCIATE LEARNING[1]

JOHN OLIVER COOK AND MORTON EDWARD SPITZER

North Carolina State College

Cook and Kendler (1956) and Cook (1958) have shown that Prompting (stimulus term, then response term, then overt practice of the response term) is superior to Confirmation (stimulus term, overt practice of the response term, and then the response term) in learning. The present experiment was undertaken to discover some reasons for this superiority. On the basis of a theoretical model (Cook & Kendler, 1956) plus some conditioning principles, it was hypothesized that the superiority of Prompting was due, at least in part, to a shorter between-term delay, the *omission* of overt practice of the response, or both. Since the temporal interval between the presentations of the two terms of a pair is shorter in Prompting (.25 sec.) than in Confirmation (3.25 sec.), and since the longer the S-R delay is, the poorer the learning will be, it was predicted that a short delay would yield a better learning performance than a long delay. Moreover, on the basis of the theoretical model, it was believed that the occurrence of overt practice during the delay interval would interfere with the establishment of the essential S-R connection that must take place over the span of the interval, and might also interfere with the learning of the response term as such.

Using the same materials as in the previous experiments, the hypotheses were tested by means of a 2 × 2 design consisting of the following four conditions (A) Prompting—No Overt Practice, (B) Confirmation—No Overt Practice, (C) Prompting—Overt Practice, and (D) Confirmation—Overt Practice. (The "program" in Cond. D is essentially the same as the program employed in Skinner's (1958) teaching machines.) There were 35 male undergraduates per condition. The *S*s were run to mastery with a test after every three of the 27 training trials.

As shown in Tables 1 and 2, the results, in terms of the number of correct items (the *C* score), bear out the predictions that Cond. A will be best (because of a short S-R delay and no intervening practice to interfere) and that Cond. D will be worst (because of a long S-R delay and because of the interfering effect of overt practice during the S-R delay interval). The relative positions of Cond. B and C were not predicted because the

[1] This research was supported by a grant from the Faculty Research and Professional Development Fund of North Carolina State College.

TABLE 1

MEANS IN SUCCESSIVE TESTS

Tests	*C* Scores				*D* Scores			
	A	B	C	D	A	B	C	D
1	2.20	1.69	1.60	0.74	3.23	3.34	3.63	2.57
2	5.14	3.80	3.43	1.86	5.69	4.77	5.00	3.77
3	7.14	6.54	5.71	3.80	7.63	7.26	6.51	5.00
4	8.77	8.06	7.57	5.06	8.80	8.40	7.83	6.17
5	9.57	8.80	8.57	6.57	9.57	8.97	8.80	7.17
6	9.51	9.26	9.09	7.57	9.51	9.40	9.17	8.06
7	9.57	9.51	9.54	8.29	9.63	9.57	9.60	8.69
8	9.86	9.71	9.77	8.63	9.86	9.71	9.77	9.00
9	9.94	9.71	9.80	8.71	9.97	9.71	9.80	9.03

TABLE 2

SIGNIFICANCES OF DIFFERENCES AMONG MEANS (MANN-WHITNEY *U* TEST)

Tests	*C* Scores						*D* Scores					
	A > B	C > D	A > C	B > D	A + B > C + D	A + C > B + D	A > B	C > D	A > C	B > D	A + B > C + D	A + C > B + D
1	.067	.001	.007	.001	.000	.001	.295	.000	.236	.000	.134	.000
2	.010	.002	.003	.000	.000	.001	.041	.000	.119	.000	.013	.000
3	.142	.002	.020	.000	.000	.004	.166	.002	.017	.000	.000	.007
4	.113	.000	.011	.000	.000	.001	.192	.002	.014	.000	.000	.007
5	.026	.002	.017	.001	.000	.001	.029	.002	.022	.001	.000	.001
6	.359	.015	.171	.005	.005	.033	.356	.021	.179	.005	.006	.048
7	—	.055	—	.031	.078	.462	—	.048	—	.029	.085	—
8	—	.020	—	.054	.092	—	—	.020	—	.055	—	—
9	—	.192	—	—	—	—	—	—	—	—	—	—

relative importance of the two variables was not known.

The two variables also proved to have a deleterious effect (Tables 1 and 2) upon the number of different legitimate responses—excluding duplicates—that a subject produces on a test (*D* score). A legitimate response is any one of the response terms used in the experiment, whether it is elicited by its proper stimulus or by some other one.

The conclusions are that: (*a*) Overt practice interferes both with learning the response term as such and with connecting it with its proper stimulus; (*b*) S-R delay interferes with the process of connecting a response to its proper stimulus, but it has no very consistently reliable effect upon learning the response term as such.

References

Cook, J. O., & Kendler, T. S. A theoretical model to explain some paired-associate learning data. In G. Finch & F. Cameron (Eds.), *Symposium on Air Force human engineering, personnel, and training research.* Washington, D. C.: Nat. Acad. Sci.—Nat. Res. Coun., Publ. 455, 1956. Pp. 90–98.

Cook, J. O. Supplementary report: Processes underlying learning a single paired-associate item. *J. exp. Psychol.*, 1958, **56**, 455.

Skinner, B. F. Teaching machines. *Science*, 1958, **128**, 969–977.

(Received May 31, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

SUPPLEMENTARY REPORT: CONTEXT EFFECTS ON ABSOLUTE JUDGMENTS OF LENGTH [1]

CARYL-ANN MILLER AND TRYGG ENGEN

Brown University

The present study is an extension of that reported by Campbell, Lewis, and Hunt (1958) on pitch. It was assumed that a further reduction in the arbitrariness of the response systems could be obtained by employing absolute judgment of line length rather than pitch.

Preliminary experiments with 99 *S*s showed that both drawing and estimating by eye yield reliable and reasonably accurate mean judgments of line lengths from 1.00 to 15.50 in. both for the very first (context free) line presented as well as for each of several lines presented in random sequences. There was only a tendency for overestimation by eye and underestimation by drawing (see also Reese et al. [1953]).

In the main experiment a systematic presentation of stimuli was used, following the procedure of Campbell et al., except that *S* made a total of 75 judgments (rather than 130), 25 in each phase and five per stimulus length in each phase. Line lengths of .75, 1.50, 3.25, 4.50, 6.75, 8.25, 10.75, 13.50, and 15.25 in. were drawn in the center of 22 × 28-in. cardboards. Extraneous cues about repetition of stimuli were eliminated by presenting a different cardboard (75 in all) for each judgment. Forty adult *S*s who were naive about the purpose of the experiment and had no special experience in estimating length were divided into two groups of 20, one (short–long group) judging the lines in the Phase 1, 2, 3 order, the other (long–short group) judging the lines in the reverse Phase 3, 2, 1 order. The lines shown in Phase 1 ranged from .75 to 6.75 in., the length being shifted gradually during Phase 2 until the range was 6.75 to 15.25 in Phase 3. The *S* was instructed to estimate by eye and report to *E* the length of each line independently and as accurately as possible to the nearest fraction of an inch.

The main results are presented in Fig. 1 where mean judgments per trial are plotted for the 6.75-in. line, the only common reference point for all three phases. Analysis of variance (Type I, Lindquist, 1953) indicates that only the Group by Trial interaction is reliable ($P < .001$, $F = 26.98$, 2 and 76 df). Thus context effects are indicated because judgments of the 6.75-in. line vary according to the stimulus context (or phase) within which the stimulus is presented. Differences between groups and trials are not reliable. Context effects on the other lines were in the same direction, those lengths closest to the 6.75-in. line being affected the most. As in the preliminary experiments, comparison of mean judgments and actual line lengths (.75—15.25 in.) indicated no appreciable loss in accuracy, although there

[1] This report is based on an Honors thesis written by the first author under the direction of the second author.

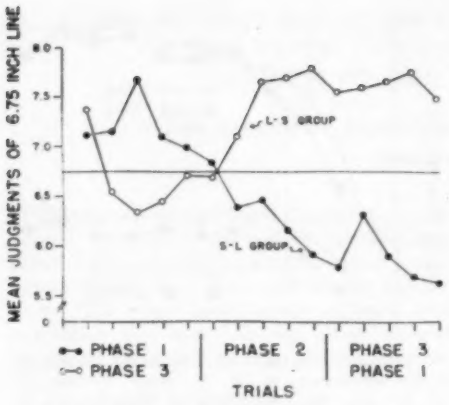


FIG. 1. Mean judgments of the 6.75 line in 15 trials distributed in three phases.

appears to be a tendency for *S*s who start with short lines to underestimate and *S*s who start with long lines to overestimate actual line length.

These results agree with those reported by Campbell et al., where comparison is feasible, in showing evidence of context unconfounded by the semantics of the response system. It should be noted that the curves obtained here cross approximately at the reference point, while in the cited pitch experiment the crossing was somewhat below the common reference point, indicating a constant error in judgment. The further minimization of arbitrariness of the response system in the present experiment appears to have improved the clarity of the results.

REFERENCES

CAMPBELL, T. C., LEWIS, N. A., & HUNT, W. A. Context effects with judgmental language that is absolute, extensive, and extra-experimentally anchored. *J. exp. Psychol.*, 1958, **55**, 220–228.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.

REESE, E. P., REESE, T. W., VOLKMANN, J., & CORBIN, H. H. (Eds.), Psychophysical research; summary report 1946–1952. *USN, Spec. Dev. Cent. Tech. Rep. 1953, SDC-131-1-5.*

(Received June 29, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

SUPPLEMENTARY REPORT: RUNWAY PERFORMANCE AS A FUNCTION OF REINFORCEMENT SCHEDULE AND ALLEY LENGTH[1]

EDWARD L. WIKE AND ROBERT REMPLE

University of Kansas

Studies of intermittent reward upon acquisition have yielded conflicting results. A survey of recent studies suggested that 100%-*S*s ran slower than 50%-*S*s in short (13–36 in.) alleys (e.g., Goodrich, 1959), while they were equal to or faster than 50%-*S*s in long (48–84 in.) alleys (e.g., Lewis & Cotton, 1957). The present problem was to study acquisition under 50% and 100% reward in short (SH) and long (L) runways. Since alley length was varied, the results also had bearing upon within-chain delay (Besch & Reynolds, 1958).

Method. The *S*s were 48 naive Sprague-Dawley male albino rats, 90–104 days old. The apparatus consisted of a starting section, two alleys (1 ft. and 5 ft.), and a goalbox. The starting section included an 8-in. starting box, a 14-in. alley, and a hurdle, ½ in. wide and 1 in. high, which was fastened to the floor 5 in. from the starting door. Raising the door started one timer; *S*'s breaking a light beam 12 in. from the door started another. A 13.5-in. × 8-in. goalbox joined the alley to form an L. When *S* entered the goal, *E* stopped the timers. The maze was 4-in. wide inside, with 6-in. walls; it was painted gray and covered with mesh.

On Days 1–7 *S*s were gentled and their home feeding of mash was reduced by 15-min. intervals from 1 hr. to 30 min. On Day 8, 6 to 7 *S*s explored a runway for 5 min; half were randomly assigned to Alley L and half to SH. Next, pairs of *S*s ate eight .10-gm. food pellets in the goalbox, and 15 min. later were fed for 30 min. in the home cage. On Day 9, pairs of *S*s explored the maze for 3 min. On Day 10, each *S* ate four pellets in the goalbox.

On Days 11 and 12 each *S* had one trial in the full maze. Thereafter *S* had two trials per day, spaced by 5 to 10 min. From each runway group 12 *S*s were randomly assigned

[1] Supported by a research grant (M-1967) from the National Institute of Mental Health, United States Public Health Service, and by the General Research Fund of the University of Kansas.

to 100% reward and 12 *Ss* to 50% (24-trial pattern = RNRNNRNRRNNRRRNNNR-RNNNRR). On R trials *S* had two pellets in the reward dish; on N trials the dish was empty and *S* was confined for 35 sec. Training continued for 44 trials.

Results and discussion. Starting times (from start to beam, Fig. 1) and running times (from beam to goalbox, Fig. 2) were transformed into speeds (feet per second). An analysis of variance of starting speeds revealed that Blocks of Trials and Runways × Blocks were significant ($F = 190.65$, $df = 10/440$, $P < .001$, and $F = 2.70$, $df = 10/440$, $P < .001$, respectively.) It is apparent (Fig. 2) that *Ss* ran *faster* in Alley L than in SH. Runways, Blocks, and Runways × Blocks were significant ($F = 182.38$, $df = 1/44$, $P < .001$, $F = 92.29$, $df = 10/440$, $P < .001$, and $F = 12.05$, $df = 10/440$, $P < .001$, respectively). Plots of the means for first daily trials disclosed a similar pattern of results. It was expected that the 50%-*Ss* would outperform the 100%-*Ss* late in training in Alley SH, and be equal or inferior in Alley L. But the second-order interaction failed to be significant for either response measure, and reward schedule did not have any significant effect. Accordingly, the results for the 50%- and 100%-*Ss* in Alley SH were not in agreement with those reported by Goodrich (1959).

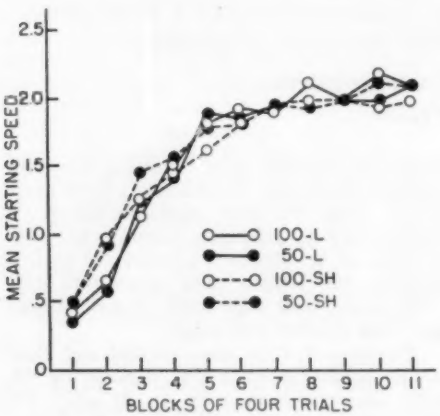


FIG. 1. Mean starting speeds.

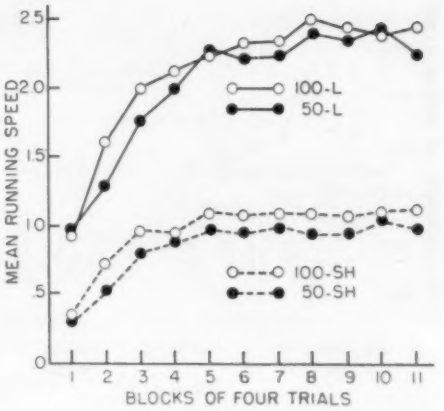


FIG. 2. Mean running speeds.

Spence's theory (1956) would predict faster starting speeds on a short runway than on a long runway. The only evidence for this prediction consisted of a significant Runway × Blocks interaction which, from an inspection of Fig. 1, is seen to be due in part to more rapid starting of *Ss* in the first phase of training in Alley SH. The faster running speeds of *Ss* in Alley L may reflect the fact that *Ss* were unable to attain maximum speed in Alley SH. These results suggest the need for parametric studies of within-chain delay.

REFERENCES

BESCH, N. F., & REYNOLDS, W. F. Alley length and time of food deprivation in instrumental reward training. *J. exp. Psychol.*, 1958, **56**, 448–452.

GOODRICH, K. P. Performance in different segments of an instrumental response chain as a function of reinforcement schedule. *J. exp. Psychol.*, 1959, **57**, 57–63.

LEWIS, D. J., & COTTON, J. W. Learning and performance as a function of drive strength during acquisition and extinction. *J. exp. Psychol.*, 1957, **50**, 189–194.

SPENCE, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

(Received August 28, 1959)

Journal of Experimental Psychology
Vol. 59, No. 4, 1960

SUPPLEMENTARY REPORT: INFLUENCE OF INTERTRIAL INTERVAL DURING EXTINCTION ON SPONTANEOUS RECOVERY OF CONDITIONED EYELID RESPONSES[1]

ELLEN Y. BEEMAN, THOMAS F. HARTMAN, AND DAVID A. GRANT

University of Wisconsin

Howat and Grant (1958) showed that with spaced extinction there were significantly fewer CRs or trials required to reach a criterion of extinction than with massed extinction. The mean spontaneous recovery was essentially the same following either massed or spaced extinction at both 20-min. and 24-hr. recovery intervals. The unexpected effect of distribution of extinction on rate of extinction made it desirable to repeat the 24-hr. recovery part of the experiment using a fixed number of acquisition and extinction trials and varying distribution during acquisition.

Method. The general apparatus, procedure, and stimuli were the same as in the earlier experiment. However, all 120 *Ss* received 60 acquisition trials. For 60 *Ss* trials were spaced; for the other 60 they were massed. Thirty *Ss* in each of these groups immediately received 15 massed extinction trials, and the other 30 received spaced extinction. Approximately 24 hr. later all *Ss* received 15 spaced extinction trials as a test of spontaneous recovery.

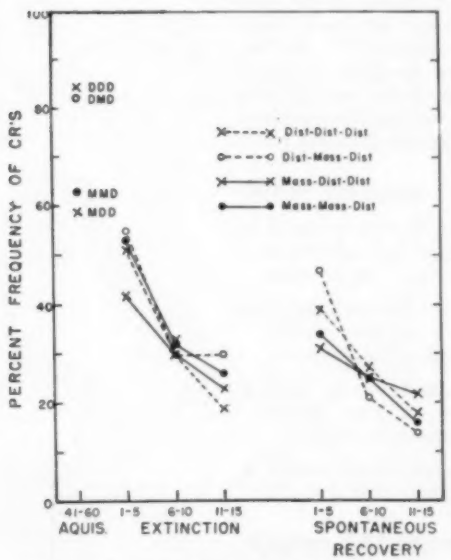


FIG. 1. Percentage frequency of CRs for all groups during the last 20 acquisition trials and on successive 5-trial blocks during extinction and spontaneous recovery test.

TABLE 1

ANALYSES OF VARIANCE

Source	*df*	*F*	
		Extinction	Spontaneous Recovery
Dist. Acq.	1	0.470	1.176
Dist. Ext.	1	1.978	0.077
Interaction	1	0.541	0.045
Error (*MS*)	116	(332.63)	(387.85)

Results. Figure 1 shows the percentage of CRs on the last 20 trials of acquisition and on 5-trial blocks during extinction and spontaneous recovery tests. The effects of distribution on acquisition were conventional and were statistically highly significant.

Table 1 gives the summary of the analysis of variance of the over-all percentage frequency of CRs during extinction, Column A, and the spontaneous recovery test series, Column B, after these percentages had been subjected to the arcsin transformation. Distribution during acquisition and extinction had no statistically significant effect upon percentage CRs during extinction or on percentage CRs during the spontaneous recovery test trials. Although there is some indication in Fig. 1 of greater spontaneous recovery for the groups receiving distributed acquisition, this result was not statistically significant.

In view of the very large number of cases and the high power of the resulting *F* tests, it is safe to conclude that distribution during extinction has little effect upon initial extinction nor upon spontaneous recovery. In the first respect the results of the earlier experiment are not confirmed. The present experiment agrees entirely with the first experiment in finding no important effect of the distribution of original extinction upon 24-hr. spontaneous recovery. Although ini-

[1] This research was supported in part by the National Science Foundation and in part by the Research Committee of the Graduate School of the University of Wisconsin with funds provided by the Wisconsin Alumni Research Foundation.

tial spontaneous recovery trials favored the massed extinction groups as before, the results for the massed and distributed groups were essentially identical on the total spontaneous recovery test, and the very sensitive *F* test was not significant.

It is not clear precisely what is the source of the difference in results of the two experiments, but it might be noted that in spite of a smaller number of *Ss* in the earlier experiment, using trials to a criterion and CRs to a criterion during acquisition, extinction, and spontaneous recovery testing resulted in smaller error variances so that the earlier experiment was actually more sensitive than the present study.

References

HOWAT, M. G., & GRANT, D. A. Influence of intertrial interval during extinction on spontaneous recovery of conditioned eyelid responses. *J. exp. Psychol.*, 1958, 56, 11–15.

(Received August 7, 1959)

Wiley BOOKS

Ready in August—

EXPERIMENTAL PSYCHOLOGY

By Burton G. Andreas,
Associate Professor of Psychology,
University of Rochester.

A new text which suggests possible directions for future research . . .

Andreas' *Experimental Psychology* not only introduces the student to methods, techniques, and important research findings, it also gives him an appreciation of the interest—and indeed—fascination involved in building a science of behavior. The book does this by presenting the subject as an active ongoing enterprise—by showing both what has been accomplished and what needs to be accomplished in the future. Throughout the book, unfinished business and methodological pitfalls are cited and discussed, and further experiments which ought to be done are mentioned.

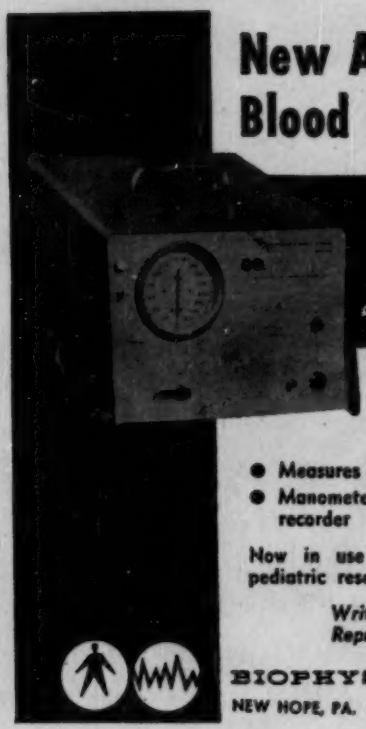
Other features . . .

- Devotes two chapters to statistical methods and techniques in experimentation
- Describes and refers to experiments which can be easily modified for duplication in the laboratory
- Can be easily adapted to meet the needs and requirements of individual instructors
- Discusses the relation of theory and empirical investigation

Approx. 544 pages. Prob. $6.95.

Watch for more features of Andreas' *Experimental Psychology* in the next issue of this journal

JOHN WILEY & SONS, Inc.
440 Park Avenue South,
New York 16, N.Y.



New Advance in Blood Pressure Measurement

CONTINUOUS SYSTOLIC MONITOR

- Accurate *continuous* measurement of blood pressure
- No arterial puncture—adult, pediatric and animal application
- Measures response to emotional and external stimuli
- Manometer pressure gauge plus connection for auxiliary recorder

Now in use in pharmacology laboratories, psychological and pediatric research, operating and recovery rooms.

Write for Brochure CSM-4 or Demonstration Representatives throughout U.S. and Canada

BIOPHYSICAL ELECTRONICS, INC.
NEW HOPE, PA.